

535,407

Rec'd PCT/PTO 18 MAY 2005

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2004 年 6 月 3 日 (03.06.2004)

PCT

(10) 国際公開番号
WO 2004/046997 A1(51) 国際特許分類⁷: G06F 19/00, 17/60, 17/30

(21) 国際出願番号: PCT/JP2003/014733

(22) 国際出願日: 2003 年 11 月 19 日 (19.11.2003)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2002-336916
2002 年 11 月 20 日 (20.11.2002) JP

(71) 出願人 (米国を除く全ての指定国について): 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒101-8010 東京都千代田区神田駿河台四丁目6番地 Tokyo (JP).

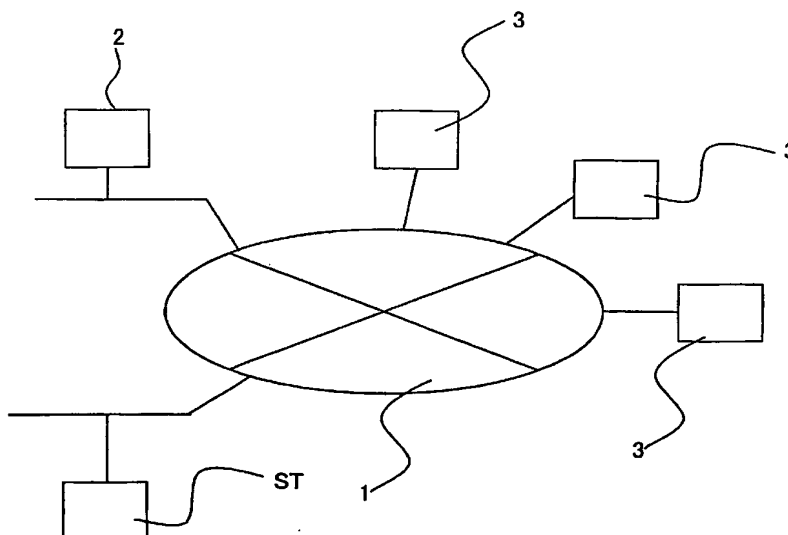
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 加藤 孝昌 (KATO, Takamasa) [JP/JP]; 〒101-8010 東京都千代田区神田駿河台四丁目6番地 株式会社日立製作所内 Tokyo (JP). 森本 健郎 (MORIMOTO, Takeo) [JP/JP]; 〒101-8010 東京都千代田区神田駿河台四丁目6番地 株式会社日立製作所内 Tokyo (JP). 松尾 仁司 (MATSUO, Hitoshi) [JP/JP]; 〒185-8601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内 Tokyo (JP). 伴 秀行 (BAN, Hideyuki) [JP/JP]; 〒185-8601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内 Tokyo (JP). 久光 徹 (HISAMITSU, Toru) [JP/JP]; 〒185-8601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内 Tokyo (JP). 神山 卓也 (KAMIYAMA, Takuya) [JP/JP]; 〒185-8601 東京都国

[続葉有]

(54) Title: DATA PROCESSING SYSTEM USING BASE SEQUENCE-RELATING DATA

(54) 発明の名称: 塩基配列関連情報を用いた情報処理システム



(57) Abstract: By effectively using differences in base sequence data among individuals, a system for providing semantic data beneficial to each individual and/or data relating to the semantic data is constructed. The above system has the step a of receiving base sequence-relating data of a definite individual and the step b of specifying a base sequence-relating data group including the base sequence-relating data agreeing with the base sequence-relating data received above from a memory device in which base sequence-relating data groups including plural sets having positional data showing a position in a base sequence and base sequence-relating data corresponding to the positional data connected with each other are stored for each individual.

(57) 要約: 個体間における塩基配列情報の相違を有効に利用して各個体にとって有益な意味情報及び/又は当該意味情報に関連する情報を提供するシステム

[続葉有]

WO 2004/046997 A1



分寺市 東恋ヶ窪一丁目280番地 株式会社日立製作所
中央研究所内 Tokyo (JP).

(74) 代理人: 平木 祐輔 (HIRAKI, Yusuke); 〒105-0001 東
京都 港区 虎ノ門一丁目17番1号 虎ノ門5森ビル 3階
Tokyo (JP).

(81) 指定国 (国内): CN, JP, KR, US.

(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY,
CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC,
NL, PT, RO, SE, SI, SK, TR).

添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される
各PCTガゼットの巻頭に掲載されている「コードと略語
のガイダンスノート」を参照。

を構築する。 所定の個体に関する塩基配列関連情報を受け取るステップaと、塩基配列における位置を意味する
位置情報及び当該位置情報に対応する塩基配列関連情報が互いに関連付けられたセットが複数含まれた塩基配列関
連情報群が個体毎に格納された記憶装置から、前記受け取った塩基配列関連情報と一致性がある塩基配列関連情報
を含んだ塩基配列関連情報群を特定するステップbとを有する。

明 細 書

塩基配列関連情報を用いた情報処理システム

5 技術分野

本発明は、例えば通信回線網を介して情報を提供する情報処理システムに関する。

背景技術

現在、ヒトを始めとする各種生物のゲノム塩基配列が急速に決定されており、
10 様々なデータベースにゲノム塩基配列情報が蓄積されている。例えば、インターネット等の情報通信網を介して、各種研究機関や研究者がデータベースに蓄積されたゲノム塩基配列情報を利用できるようなシステムの構築がなされつつある。

同時に、このようなゲノム塩基配列情報に含まれる塩基配列を用いて、ゲノム創薬の研究や遺伝情報の解析等が盛んに行われており、一塩基多型に代表されるよう
15 な個体間における塩基配列の相違が注目されている。一般に、個体間における塩基配列の相違とは、所定の塩基の相違が個体種中 1 % 以上の頻度で存在すると定義される多型と、所定の塩基の相違が個体種中 1 % 未満であるバリエーションとを意味している。特に、多型には、個体間における 1 個の塩基の相違である一塩基多型 (SNP ; Single Nucleotide Polymorphism) 、 1 から数十塩基 (数千塩基の場合もある)
20 が欠失又は挿入している挿入/欠失多型、 2 から数十塩基を 1 単位とする配列の繰り返し回数が相違する VNTR (Variable Number of Tandem Repeat) やマイクロサテライト多型 (繰り返し配列が 2 ~ 4 塩基程度のもの) が知られている。

このような多型は、個体間におけるタンパク質のアミノ酸配列の相違や、個体間における所定の遺伝子に関する発現効率の相違等に影響を及ぼすことがある。この
25 ような影響により、例えば、所定の疾病に対する罹患可能性が個体間で異なったり、所定の薬剤に対する感受性が個体間で異なることが知られている。

ところが、多型等の個体間における塩基配列情報の相違を有効に利用して、各個体にとって有益な意味情報を提供するようなシステムは構築されていないのが現状である。

発明の開示

そこで、本発明は、このような現状に鑑み、個体間における塩基配列情報の相違を有効に利用して各個体にとって有益な意味情報及び/又は当該意味情報に関連する

5 情報を提供するシステムを構築することを目的とする。

上述した目的を達成した本発明に係る塩基配列に関する情報処理方法は、所定の個体に関する塩基配列関連情報を受け取った後、個体毎に塩基配列関連情報群を有する記憶装置から前記受け取った塩基配列関連情報と一致性がある塩基配列関連情報を含んだ塩基配列関連情報群を特定する方法である。本方法によれば、塩基配列

10 関連情報群を特定することによって、前記受け取った塩基配列関連情報が帰属する個体を特定することもまたできる。

また、本発明に係る塩基配列に関する情報処理方法は、所定の個体に関する個体関連情報を受け取り、特定した塩基配列関連情報群若しくは個体と、前記受け取った個体関連情報とを関連付けて記憶することで、当該個体関連情報と個体に関する

15 塩基配列関連情報群とを関連付けることができる。

なお、本発明に係る塩基配列に関する情報処理方法において、塩基配列関連情報群を特定する際には、例えば、予め決まった位置情報に対応する塩基配列関連情報を受け取り、当該受け取った塩基配列関連情報を用いることができる。また、本発明に係る塩基配列に関する情報処理方法においては、所定の個体に関する塩基配列

20 関連情報の受け取りと、当該受け取った塩基配列関連情報と一致性がある塩基配列関連情報を含んだ塩基配列関連情報群の特定とを、唯一の塩基配列関連情報群を特定すまで繰り返すこともできる。

さらに、本発明に係る塩基配列に関する情報処理方法は、所定の個体に関して個体関連情報を受け取ることによって、当該所定の個体に関する塩基配列関連情報群

25 と個体関連情報とを関連付けたデータベースを構築することもできる。

本発明に係る塩基配列に関する情報処理方法は、前記データベースに含まれる複数の個体に関する複数の個体関連情報及び複数の個体に関する塩基配列関連情報群を統計的に処理することで、塩基配列関連情報を意味づける意味情報及び/又は当該意味情報に関連する情報を創出することができる。このとき、当該複数の個体に関

する塩基配列関連情報群は、前記データベースに含まれていても、含まれていなくても良い。また、統計的に処理した結果と複数の個体に関する複数の個体関連情報とを用いて更に統計的に処理することによって、意味情報及び/又は当該意味情報に関連する情報を創出してもよい。

- 5 一方、本発明に係る塩基配列に関する情報処理方法は、創出された塩基配列関連情報を意味づける意味情報及び/又は当該意味情報に関連する情報を受け取り、受け取った意味情報及び/又は当該意味情報に関連する情報を用いて、物品及び/又はサービスの要求情報に応じた意味情報及び/又は当該意味情報に関連する情報を提供するための記憶装置の記憶内容を構築することもできる。
- 10 なお、本発明に係る塩基配列に関する情報処理方法は、制御装置、送受信装置及び記憶装置等のハードウェアを備えるコンピュータに、各ステップを実行させるプログラムとして実現することができる。また、本発明に係る塩基配列に関する情報処理方法は、制御装置、送受信装置及び記憶装置等のハードウェアを備えるコンピュータに、各ステップを実行させるプログラムを記録した記録媒体として実現する
- 15 こともできる。さらに、本発明に係る塩基配列に関する情報処理方法は、各ステップを実行する制御装置、送受信装置及び記憶装置等のハードウェアを備える情報処理装置として実現することもできる。

その他、本発明は、請求項各項に記載されている通りの構成を有するものである。

- 本明細書は本願の優先権の基礎である日本国特許出願2002-336916号の明細書および/または図面に記載される内容を包含する。
- 20

図面の簡単な説明

図1は、本発明を適用した情報処理システムの構成を概略的に示す概略構成図である。

- 25 図2は、共用コンピュータの構成を概略的に示す概略構成図である。

図3は、メインDBに記録されたデータの一例を示す構成図である。

図4は、統計解析用コンピュータの構成を概略的に示す概略構成図である。

図5は、ゲノム関連情報DBに記録されたデータの一例を示す構成図である。

図6は、個体関連情報DBに記録されたデータの一例を示す構成図である。

図 7 は、個人用コンピュータの構成を概略的に示す概略構成図である。

図 8 は、ゲノム関連情報記録媒体に記録されたデータの一例を示す構成図である。

図 9 は、個体関連情報とゲノム関連情報とを関連連付ける情報処理システムにおいて、統計解析用コンピュータ及び個人用コンピュータでの処理を示すフローチャートである。

図 10 は、個体関連情報収集画面の一例として示す画面イメージである。

図 11 は、ゲノム関連情報 DB の全データから、多型番地における各多型パターンの出現頻度を算出したデータの一例を示す構成図である。

図 12 は、ゲノム関連情報 DB と個体関連情報 DB とを用いて、個体関連情報について多型番地毎の各多型パターンの出現頻度を算出したデータの一例を示す構成図である。

図 13 は、図 11 及び図 12 に示したデータから、個体関連情報について多型番地毎の各多型パターンの割合値を算出したデータの一例を示す構成図である。

図 14 は、図 13 に示したデータから、個体関連情報について多型番地毎の各多型パターンの相対値を算出したデータの一例を示す構成図である。

図 15 は、所定の疾病に対する罹患可能性を提供するシステムにおいて、共用コンピュータ及び個人用コンピュータでの処理を示すフローチャートである。

図 16 は、図 15 の続きであり、所定の疾病に対する罹患可能性を提供するシステムにおいて、共用コンピュータ及び個人用コンピュータでの処理を示すフローチャートである。

図 17 は、所定の疾病に対する罹患可能性を提供するシステムにおいて、共用コンピュータ及び個人用コンピュータでの他の処理を示すフローチャートである。

図 18 は、所定の疾病に対する罹患可能性を提供するシステムにおいて、共用コンピュータ及び個人用コンピュータでの更に他の処理を示すフローチャートである。

符号の説明

1…通信回線網、2…共用コンピュータ、3…個人用コンピュータ、ST…統計解析用コンピュータ

発明を実施するための最良の形態

以下、図面を参照して本発明を詳細に説明する。

本発明を適用した実施の形態として、利用者に対して所定の疾病の罹患可能性等の意味情報を提供する情報処理システムにおいて、当該意味情報を創出するシステムについて説明する。すなわち、情報処理システムは、利用者に対して提供する罹患可能性等の意味情報を、個体に関する情報（以下、個体関連情報）から創出するシステムである。

以下の説明においては、個体関連情報から意味情報を創出する情報処理システムについて説明するが、説明の都合上、簡略化したモデルとして説明する。情報処理システムは、図 1 に示すように、インターネット等の通信回線網 1 と、通信回線網 1 に接続された共用コンピュータ 2 と、通信回線網 1 に接続された少なくとも 1 以上の個人用コンピュータ 3 と、通信回線網 1 に接続された統計解析用コンピュータ S T とを備え、通信回線網 1 を介して共用コンピュータ 2 と個人用コンピュータ 3 と統計解析用コンピュータ S T との間のデータ通信を可能としている。

共用コンピュータ 2 は、図 2 に示すように、当該共用コンピュータ 2 の動作を全て制御する CPU 4 と、情報及びプログラムの実行指示等を入力できるキーボード及びマウス等の入力装置 5 と、ディスプレイ装置等の表示装置 6 と、一時的な情報及び書き換え不可能な情報等が記録されるメモリー 7 と、各種データを格納しているデータベース 8 と、これらメモリー 7 及びデータベース 8 に対して所定の情報を書き込む記録装置 9 と、通信回線網 1 を介して個人用コンピュータ 3 との間で情報の送受信を行う送受信装置 17 とから構成されている。

共用コンピュータ 2 におけるメモリー 7 は、それぞれ異なる種類の情報を記録するメモリー部 A 10 及びメモリー部 B 11 と、例えば個人用コンピュータ 3 や表示装置 6 に表示させる画像データを記録した画面メモリー 12 と、本システムを動作させるための処理プログラム 13 とから構成されている。なお、共用コンピュータ 2 においては、画面メモリー 12 及び処理プログラム 13 等を内部のメモリー 7 に有さず、通信回線網 1 を介して共用コンピュータ 2 と接続された外部記憶装置（図示せず）に有するものであってもよい。

共用コンピュータ 2 におけるデータベース 8（記憶装置）は、多型番地、多型パ

ターン及び意味情報が記録されたメインDB 1 4 と、メモリー部 A 1 0 に記録された情報を保存する保管用DB-A 1 5 と、メモリー部 B 1 1 に記録された情報を保存する保管用DB-B 1 6 とから構成されている。メインDB 1 4 は、図 3 に示すように、多型番地と、当該多型番地で取りうる複数の多型パターンと、当該複数の多型パターンそれぞれを意味づける意味情報とが関連付けられて記録されている。また、メインDB 1 4 には、複数の多型番地における多型パターンの組合せ（例えば、ハプロタイプ）を意味づける意味情報が記録されていても良い。

ここで、「多型番地（位置情報）」とは、少なくとも、塩基配列における多型が存在する位置を意味する。なお、一般的に多型とは、例えば、いわゆるSNP(single nucleotide polymorphism)、RFLP(restriction fragment length of polymorphism)、VNTR(variable number of tandem repeat)、マイクロサテライト等を含んでいる。しかし、本明細書において使用する「多型」は、これらに限定されず、個体種中 1 %未満の頻度でしか存在しない塩基及び塩基配列の変化（バリエーション）も含む意味とする。したがって、「多型番地」は、個体種中 1 %未満の頻度でしか存在しない塩基及び塩基配列の変化を示す、塩基配列における位置も含む意味である。すなわち、「多型番地」とは、数値、文字及び記号等を組み合わせて、多型等を示す位置を表すものである。多型番地は、特に限定されないが、例えば、染色体番号と多型が存在する遺伝子を表す記号と当該遺伝子における多型の存在位置を示す数値との組み合わせにより表記することもできるし、多型が存在する遺伝子を示す記号と当該遺伝子における多型の存在位置を示す数値との組み合わせであってもよい。

また、多型番地は、多型毎に付与される多型固有の表記であっても良い。多型番地として多型固有の表記を使用する場合、多型番地は塩基配列中の位置を直接的には示さないが、多型固有の表記に基づいて間接的に位置を知ることができる。したがって、「多型番地」は、多型固有の表記も含む意味である。

「多型パターン（塩基配列関連情報）」とは、個体間において相違する塩基配列の情報であり、少なくとも、多型における塩基又は塩基配列のパターンを含む意味である。さらに「多型パターン」は、多型に限らず、個体種中 1 %未満の頻度でしか存在しない塩基及び塩基配列のパターンも含む意味である。

例えば、A又はGを取ることが知られている多型番地において、「多型パター

- ン」は、「A」及び「G」のいずれかで表される。また、「多型パターン」は、相同染色体におけるヘテロ接合体又はホモ接合体を示すものであってもよい。この場合、例えば、A又はGを取ることが知られている多型番地において、「多型パターン」は、「AA」、「GG」及び「AG」のいずれかで表現できる。さらに、「多型パターン」は、所定の多型番地を取りうるパターンを直接的に表記するものではなく、間接的に表記するものであっても良い。すなわち、「多型パターン」は、例えば、A又はGを取ることが知られている多型番地において「A」を取る場合に「アレル1」とし、「G」を取る場合に「アレル2」と表記してもよい。また、「多型パターン」が上述したように「AA」、「GG」及び「AG」のいずれかで表現できる場合、例えば、「AA」で表現できるときに「 α 」、「GG」で表現できるときに「 β 」、「AG」で表現できるときに「 γ 」と表記してもよい。その他「多型パターン」の表記例としては、多型がマイクロサテライトの場合には「繰り返し数」を表す数値で、多型が挿入、欠失型の場合には「有/無」を表す記号で表記してもよい。
- 15 また更に、各多型番地における「多型パターン」は、所定の規則や取り決めに従って、例えば、「多型1」、「多型2」、「多型3」等と表記されても良い。例えば、各多型番地において、「多型パターン」がとり得る頻度の高い順に、「多型1」、「多型2」、「多型3」と表記できる。この場合、例えば、各多型番地におけるそれぞれの「多型1」は必ずしも同じ内容を表すものではない。すなわち、例えば、ある多型番地の「多型1」は最もとり得る頻度が高い「AA」を表し、別の多型番地「多型1」は最もとり得る頻度が高い「GG」を表すことになる。なお、本実施の形態においては、「多型パターン」を「多型1」、「多型2」及び「多型3」等のいずれかで表記する方法を使用している。ところで、本システムにおいては、多型パターンは、暗号化されていても暗号化されていなくても差し支えない。
- 25 ここで、「意味情報」とは、「多型パターン」に関連づけられた情報であり、例えば、薬剤に対する応答性、薬剤に対する副作用、疾患及び障害に対するリスク、体質・性質、体質・性質等に基づく生活習慣アドバイス、タンパク質相互作用など、「多型パターン」の相違に起因する様々な情報を意味する。

なお、「意味情報」としては、「多型パターン」の相違に起因する様々な情報を

直接表しても良く、また、当該情報を意味する記号などを用いて間接的に表しても良い。「意味情報」は、ゲノム・遺伝子に関する研究が進むことにより種類が増加するとともに訂正が行われる種類の情報であり、常にバージョンアップすることが好ましい。すなわち、「意味情報」は、ゲノム・遺伝子の研究成果を用いてデータベースを更新することによって、蓄積量が増加・減少してより精度の高いものとなる。

なお、直接「多型パターン」には関連づけられていないが「意味情報」から更に導き出される情報は、「意味情報に関連する情報」である。「意味情報」が「疾患に対するリスク」である場合、当該リスクがある一定の水準を超えたときに、例えば特定の「健康診断検査項目」が導き出される。この特定の「健康診断検査項目」が「意味情報に関連する情報」である。

本実施の形態において意味情報は、図3に示すように、少なくとも、所定の「多型番地」及び「多型パターン」に関連づけられた「多型パターンに対する注釈情報」としてメインDB14に記録されている。また、意味情報には、所定の「多型番地」に対応する「多型分類」及び「分類（疾患名）」等が関連づけられている。すなわち、所定の「多型番地」が所定の「多型パターン」である場合、疾患名の種類と当該疾患に対する罹患可能性を示す注釈情報（意味情報）を得ることができる。したがって、例えば、意味情報は、複数の多型番地に対応するそれぞれの多型パターンの組み合わせ（例えば、ハプロタイプ）に対して関連付けることもできる。すなわち、複数の多型番地における多型パターンの組み合わせ毎に、所定の疾患に対する異なる罹患可能性を示す注釈情報（意味情報）を関連付けることができる。この場合、複数の多型番地が所定の多型パターンの組み合わせである場合、所定の疾患に対する罹患可能性を示す注釈情報（意味情報）を得ることができる。

また、意味情報には、所定の基準で決定した「公開レベル」を関連づけることもできる。例えば、「公開レベル」を決定する際の基準としては、意味情報、すなわちここでは「分類（疾患名）」の罹患可能性を公開することによる個人に対する不測の不利益等を考慮して定めることができる。詳細には、共用コンピュータ2において、法律、規則又は自らの行動基準若しくは利用者との契約等に鑑みて、公開することが相応しくない意味情報については、公開しないような「公開レベル」を決

定することができる。この場合、本システムでは、公開不可を意味する「公開レベル」に関連付けられた罹患可能性を示す注釈情報については、利用者に対して開示することはない。これにより、利用者に対して不測の不利益となりうる意味情報を与えることや、契約者以外に意味情報が開示されることを防止できる。

- 5 なお、利用者がインフォームドコンセント等により、所定の「公開レベル」を関連づけた意味情報の開示を容認することにより、利用者に対して、所定の「公開レベル」が関連づけられた意味情報を公開するようなシステムであってもよい。

また、「公開レベル」は、例えば「1, 2, 3, …」又は「a, b, c, …」といった3以上の複数の段階として設定することができる。この場合、共用コンピュータ2側では、利用者の年齢、資格及び利用者との契約の有無等、利用者の種類に応じてレベルを設定することができる。なお、インフォームドコンセント等によって、
10 所定の公開レベル以上（又は未満）の公開レベルに関連付けられた罹患可能性を示す注釈情報のみが利用者側に対して提供されるように、当該利用者側が公開レベルを選択することもできる。

- 15 なお、データベース8において、保管用DB-B16には、例えば、本システムを利用する要求者個人の遺伝情報である塩基配列関連情報といったデータを記録することができる。また、保管用DB-A15には、例えば、本システムを利用する要求者を特定する情報といったデータを記録することができる。このように、保管用DB-A15及び保管用DB-B16に、個人の遺伝情報と個人を特定する情報とを分けて記録する
20 ことによって、要求者の遺伝情報と、要求者を特定するデータとを関連付け難くなる。

なお、共用コンピュータ2は、データベース8を内部に有するものに限定されず、通信回線網1を介して共用コンピュータ2に接続された外部データベース（図示せず）を有するものであってもよい。また、共用コンピュータ2は、内部に複数のデータベース8を有するものであってもよいし、内部のデータベース8と通信回線網1を介して共用コンピュータ2に接続された外部データベースとを有するものであ
25 っても良い。

統計解析用コンピュータSTは、図4に示すように、当該統計解析用コンピュータSTの動作を全て制御するCPU30と、情報及びプログラムの実行指示等を入力でき

るキーボード、マウス等の入力装置 3 1 と、ディスプレイ装置等の表示装置 3 2 と、一時的な情報及び書換え可能な情報や書換え不可能な情報等が記録されるメモリー 3 3 と、各種データを格納しているデータベース 3 4 と、これらメモリー 3 3 及びデータベース 3 4 に対して所定の情報を書き込む記録装置 3 5 と、通信回線網 1 を介して
5 共用コンピュータ 2 及び個人用コンピュータ 3 との間で情報の送受信を行う送受信装置 3 6 とから構成されている。

統計解析用コンピュータ S T におけるメモリー 3 3 は、個人用コンピュータ 3 から送られた多型パターン等を一時的に記録したり、統計解析結果等を記録するメモリー部 3 7 と、例えば個人用コンピュータ 3 や共用コンピュータ 2 や表示装置 3 2
10 に表示させる画像データを記録した画面メモリー 3 8 と、本システムを動作させ、複数の個人（個体）に関する「個体関連情報」と複数の個人（個体）に関する「ゲノム関連情報」とを用いて、例えば、所定の疾患に対する罹患可能性を示す注釈情報といった意味情報を創出する処理プログラム 3 9 とが記録されている。なお、当該処理プログラム 3 9 によれば、先ず、複数の個人（個体）に関する「ゲノム関連情報」と
15 複数の個人（個体）に関する「個体関連情報」とを集積した後、当該「ゲノム関連情報」と当該「個体関連情報」とを用いて統計的に処理することで、所定の疾患に対する罹患可能性を示す注釈情報（意味情報）等を得ることができる。なお、統計解析用コンピュータ S T においては、処理プログラム 3 9 やメモリー部 3 7 を内部のメモリー 3 3 に有さず、通信回線網 1 を介して統計解析用コンピュータ S T に接続され
20 た外部記憶装置（図示せず）に有するものであっても良い。

ここで、「個体関連情報」とは、個体の性質、心理的状态、体質、体調、健康状態、病歴、生活習慣、行動・思考パターン、癖及び嗜好等、所定の個体に関するあらゆる情報が含まれる。例えば、個体関連情報としては、予め準備された質問等に対して利用者等の個人（個体）が回答することから得られる情報、或いは、例えば、利
25 用者等が医療機関や検査機関等で検査を受けることによって得られる情報を挙げる
ことができる。

また、「ゲノム関連情報」とは、所定の個体に関する複数の「多型パターン」をそれぞれ所定の「多型番地」に関連付けたデータの群（塩基配列関連情報群）である。

統計解析用コンピュータ S T におけるデータベース 3 4（記憶装置）は、複数の

個人(個体)に関する「ゲノム関連情報」を各個人(個体)毎に記録した「ゲノム関連情報DB 40」(図5に示す)を有している。なお、「ゲノム関連情報DB 40」に記録された「ゲノム関連情報」は、直接個人(個体)を特定する情報と結びつかないように匿名化されていることが望ましい。

- 5 また、データベース34は、個人用コンピュータ3より受信した「個体関連情報」を各個人(個体)毎に記録した「個体関連情報DB 41」(図6に示す)を有している。なお、「個体関連情報DB 41」に記録された「個体関連情報」は、直接個人(個体)を特定する情報と結びつかないように匿名化されていることが望ましい。

- 10 なお、統計解析用コンピュータSTは、前記データベース34を内部に有するものに限定されず、通信回線網1を介して統計解析用コンピュータSTに接続された外部データベース(図示せず)に対してアクセスするものであっても良い。

- 15 個人用コンピュータ3は、図7に示すように、当該個人用コンピュータ3の動作を全て制御するCPU20と、情報及びプログラムの実行指示等を入力できるキーボード及びマウス等の入力装置21と、ディスプレイ装置等の表示装置22と、
20 一時的な情報及び書き換え可能な情報等が記録されるメモリー23と、ゲノム関連情報記録媒体24からデータを読み取る読取り装置25と、通信回線網1を介して共用コンピュータ2との間で情報の送受信を行う送受信装置29とから構成されている。なお、個人用コンピュータ3は、通常のコンピュータに限定されず、例えば、携帯電話、個人携帯端末及びその他の移動体通信機器等、いかなる形態であつても
20 よい。

個人用コンピュータ3におけるメモリー23は、ゲノム関連情報記録媒体24からの情報等を記録するメモリー部26を有し、本情報処理システムを動作させる処理プログラム27が記録されている。

- 25 ゲノム関連情報記録媒体24には、個人のゲノム関連情報28が記録されている。ゲノム関連情報記録媒体24としては、例えば、磁気ディスクや磁気カード等の磁気記録媒体、光磁気記録方式や相変化記録方式等を適用した光学式記録媒体、半導体メモリー等を挙げることができる。また、このゲノム関連情報記録媒体24は、カード状、ディスク状、スティック状、テープ状又はドラム状等いかなる形態であつてもよい。さらに、このゲノム関連情報記録媒体24は、単一の個人(個体)の

ゲノム関連情報 28 を記録したものであってもよいが、複数の個人（個体）に関する複数のゲノム関連情報 28 を記録したものであってもよい。

ゲノム関連情報記録媒体 24 に含まれるゲノム関連情報 28 とは、少なくとも、「多型番地」及び個人（個体）の塩基配列を解析した結果として得られる所定の多
5 型番地における「多型パターン」を意味する。また、ゲノム関連情報 28 には、既往症、特徴、カルテ情報、健康診断結果といった各種情報を含んでいてもよい。

ゲノム関連情報記録媒体 24 には、ゲノム関連情報 28 として、例えば、図 8 に示すように、データ I としてゲノム関連情報 28 に固有の個別番号「Gno.」（ジナンバー）及び生年月日等の個人情報記録し、データ II として多型番地及び多型パ
10 ターンを記録し、データ III として既往症を記録し、データ IV として特徴を記録し、データ V としてカルテ情報等を記録する。すなわち、ゲノム関連情報 28 は、データ I、データ II、データ III、データ IV 及びデータ V から構成されている。データ I 及びデータ II には必須の情報が含まれており、データ III、データ IV 及びデータ V には付加的な情報から構成されている。

15 ゲノム関連情報 28 においては、塩基配列上の位置に対応する「多型番地」と、当該多型番地における「多型パターン」とをリンクさせて記録している。また、データ II には、所定の多型番地における付加的な情報を「コメント」として、「多型番地」にリンクさせて記録していてもよい。なお、データ II には、所定の個体に関する全塩基配列を記録しても良い。データ II に全塩基配列を記録した場合であっても、データ II 内に「多型番地」及び「多型パターン」が含まれることとなる。
20

なお、本発明において、個人用コンピュータ 3 及びゲノム関連情報記録媒体 24 は、それぞれ図 7 及び図 8 に示したような構成に限定されず、例えば、ゲノム関連情報記録媒体が処理プログラムを有するメモリー部を備え、個人用コンピュータが当該ゲノム関連情報記録媒体を装着して処理プログラムを動作させるような構成であ
25 ってもよい。この場合、個人用コンピュータは、ゲノム関連情報記録媒体のメモリー部に記録された処理プログラムに従って動作できる。

以上のように構成された情報処理システムにおいては、統計解析用コンピュータ ST のメモリー 33 が有する処理プログラム 39 及び個人用コンピュータ 3 のメモリー 23 に記録された処理プログラム 27 が例えば、図 9 に示すようなフローチャ

ートに従って情報処理動作する。なお、図 9 に示すフローチャートにおいて、「(統)」と記載したステップは統計解析用コンピュータ S T における処理を意味し、「(個)」と記載したステップは個人用コンピュータ 3 における処理を意味している。

- 5 本情報処理システムは、ゲノム関連情報記録媒体 2 4 を所持する各個人が個人用コンピュータ 3 を用いて通信回線網 1 を介して統計解析用コンピュータ S T にアクセスし、各個人（個体）に関する「個体関連情報」を統計解析用コンピュータ S T の「個体関連情報 DB 4 1」に、「ゲノム関連情報 DB 4 0」の中のゲノム関連情報と関連付けて登録するシステムである。なお、本情報処理システムは、複数人の
- 10 ゲノム関連情報 2 8 がそれぞれ記録されたゲノム関連情報記録媒体 2 4 を用い、各個人がゲノム関連情報記録媒体 2 4 にアクセスするようなシステムであつてもよい。

- このとき、先ず要求者は、ステップ 1 (S 1) で、通信回線網 1 を介して統計解析用コンピュータ S T にアクセスし、統計解析用コンピュータ S T に対して個体関連情報の登録を行う旨の意思表示を行う。ステップ 1 では、統計解析用コンピュータ S T が提供するウェブページにアクセスして前記意思表示を行っても良いし、統計解析用コンピュータ S T に対して電子メール等を用いてアクセスして前記意思表示を行っても良い。なお、本例においては、要求者が自らの個体関連情報を登録する場合について述べるが、これに限定されず、要求者が自分以外の個人（個体）に関する個体関連情報を登録することもできる。

- 20 次に、統計解析用コンピュータ S T は、前記意思表示を個人用コンピュータ 3 から受信した後、ステップ 2 (S 2) で、画面メモリー 3 8 から、図 1 0 に示すような「個体関連情報収集画面」を読み出し、個人用コンピュータ 3 の表示装置 2 2 に表示する。ステップ 2 では、統計解析用コンピュータ S T が提供するウェブページに個人用コンピュータ 3 がアクセスした状態で個体関連情報収集画面を表示しても
- 25 良いし、個人用コンピュータ 3 に対して送信した個体関連情報収集画面データに基づいて個人用コンピュータ 3 の表示装置 2 2 に表示しても良い。

次に、個人用コンピュータ 3 は、ステップ 3 (S 3) で、個体関連情報収集画面に従って、要求者自身の個体関連情報を入力する。すなわち、要求者は、個体関連情報収集画面に表示されている設問に対して回答し、当該個体関連情報収集画面に

回答内容を入力する。或いは、要求者は、個体関連情報収集画面に表示されている設問に対する回答を、当該個体関連情報収集画面とは異なる回答用画面に入力しても良い。また、ステップ3では、個体関連情報として利用者（要求者）等が医療機関や検査機関等で検査を受けることによって得た情報を入力しても良い。

- 5 次に、個人用コンピュータ3は、ステップ4（S4）で、個体関連情報収集画面に表示されている設問に対する回答（個体関連情報）を統計解析用コンピュータSTに対して送信する。ステップ4では、回答を入力した個体関連情報収集画面データ又は回答を入力した回答用画面データを、通信回線網1を介して送信することで前記設問に対する回答を統計解析用コンピュータSTに対して送信することができる。
- 10 また、ステップ4では、個体関連情報として利用者（要求者）等が医療機関や検査機関等で検査を受けることによって得た情報を統計解析用コンピュータSTに対して送信することもできる。

- 15 次に、統計解析用コンピュータSTは、個人用コンピュータ3から個体関連情報を受信した後、ステップ5（S5）で、個人用コンピュータ3に対して、複数の多型番地を送信する。ステップ5で送信する複数の多型番地は、既定の多型番地であっても良いし、ランダムに選択した多型番地であっても良い。

- 20 次に、個人用コンピュータ3は、統計解析用コンピュータSTから複数の多型番地を受信した後、ステップ6（S6）で、読取り装置25を駆動してゲノム関連情報記録媒体24にアクセスする。次に、個人用コンピュータ3は、ステップ7（S7）で、統計解析用コンピュータSTより受信した複数の多型番地について、それぞれ対応する多型パターンを読み出す。そして、個人用コンピュータ3は、ステップ8（S8）で、ステップ7で読み出した多型パターンを対応する多型番地と関連付けて統計解析用コンピュータSTに対して送信する。すなわち、ステップ8では、統計解析用コンピュータSTより受信した複数の多型番地それぞれについて、対応
- 25 する多型パターンを関連付けて送信する。なお、ステップ8で個人用コンピュータ3から統計解析用コンピュータSTに送信する際には、「Gno.」のような個人（個体）を特定しうる情報を送信しないことが望ましい。

ところで、本実施の形態では、ステップ5で統計解析用コンピュータSTが送信した多型番地を個人用コンピュータ3で受信し、受信した多型番地に対応する多型

パターンをステップ 8 で個人用コンピュータ 3 が統計解析用コンピュータ S T に対して送信している。しかしながら、これに限定されず、例えばステップ 4 において、個人用コンピュータ 3 が統計解析用コンピュータ S T に対して「個体関連情報」を送信するとともに、既定の多型番地と当該多型番地に対応する多型パターンとを自
5 発的に統計解析用コンピュータ S T に対して送信してもよい。この場合は、上述したステップ 5 ～ 8 までは行われず、ステップ 4 の後に以下のステップ 9 以降のステップを同様に行うことになる。

次に、統計解析用コンピュータ S T は、個人用コンピュータ 3 から多型番地及び多型パターンを受信した後、ステップ 9 (S 9) で、ゲノム関連情報 D B 4 0 にア
10 クセスする。なお、ステップ 9 では、統計解析用コンピュータ S T がゲノム関連情報 D B 4 0 を有さず、外部の機関がゲノム関連情報 D B 4 0 を有する場合には、通信回線網 1 を介して当該外部の機関が有するゲノム関連情報 D B 4 0 にアクセスする。

次に、統計解析用コンピュータ S T は、ステップ 1 0 (S 1 0) で、個人用コン
15 ピュータ 3 より受信した複数の多型番地及び多型パターンの組合せに基づいてゲノム関連情報 D B 4 0 を検索し、ゲノム関連情報 D B 4 0 に格納されたゲノム関連情報の中から、受信した複数の多型番地及び多型パターンの組合せを有する個人（個体）に関するゲノム関連情報を特定する。言い換えると、ステップ 1 0 では、ゲノム関連情報 D B 4 0 に登録された複数の個人（個体）に関する複数のゲノム関連情
20 報の中から要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報を特定する。ステップ 1 0 では、例えば、特定した要求者、すなわち個人（個体）に関するゲノム関連情報に対して「整理No.」を付与してもよい。或いは、ステップ 1 0 では、ゲノム関連情報 D B 4
0 に登録している個人（個体）に関するゲノム関連情報に対して予め付与された
25 「整理No.」を抽出しても良い。

次に、統計解析用コンピュータ S T では、ステップ 1 1 (S 1 1) で、ステップ 1 0 で特定した要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報と、個人用コンピュータ 3 より受信した個体関連情報とを関連付ける。具体的には、ステップ 1 0 で特定した要求者（要求者が送信し

た「個体関連情報」に関連する個人（個体））に関するゲノム関連情報に付与した「整理No.」、又はステップ10で特定した要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報について抽出した「整理No.」を、個体関連情報に関連付けて格納した、図6に示すような「個体関連情報DB41」を作成する。

或いは、ステップ10では、統計解析用コンピュータSTがゲノム関連情報DB40を有する場合には、ゲノム関連情報DB40に格納されている要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報に対して、個人用コンピュータ3より受信した個体関連情報を直接関連付けて格納しても良い。

以上、図9に示すフローチャートに従えば、例えば所定の要求者について、当該要求者に関する個体関連情報を、当該要求者に関するゲノム関連情報に関連付けて統計解析用コンピュータSTに登録することができる。また、複数の要求者が図9に示すフローチャートに従ってそれぞれ個体関連情報を登録することによって、統計解析用コンピュータSTは、複数の個人（個体）に関する個体関連情報を有することとなる。言い換えると、統計解析用コンピュータSTは、複数の個体に関するゲノム関連情報についてそれぞれ個体関連情報に関連付けて格納したデータベースを作成することができる。

以上のように、本システムによれば、統計解析用コンピュータSTの「ゲノム関連情報DB40」に、予め匿名化して記録しておいたゲノム関連情報に対して、後から匿名化した個体関連情報を統計解析用コンピュータSTに送信した場合であっても、当該ゲノム関連情報と当該個体関連情報とをリンクさせることができる。

なお、図9に示したフローチャートにおいては、統計解析用コンピュータSTは、ステップ5で送信した複数の多型番地について、個人用コンピュータ3より多型パターンを受信し、ステップ10でゲノム関連情報DB40を検索することによって要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報を特定している。しかしながら、本情報処理システムにおいては、要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報の特定に際してこの方法に限定されず、例えば、所定の「多型番地」及

び「多型パターン」の組合せを、順次個人用コンピュータ 3 から統計解析用コンピュータ S T に対して送信し、統計解析用コンピュータ S T が要求者の特定を行っても良い。この場合、個人用コンピュータ 3 から所定の「多型番地」及び「多型パターン」の組合せを統計解析用コンピュータ S T に対して順次自発的に送信しても良いし、統計解析用コンピュータ S T から所定の「多型番地」に対応する「多型パターン」の提出を個人用コンピュータ 3 に対して順次要求し、個人用コンピュータ 3 が当該要求に対応した「多型パターン」を「多型番地」に関連付けて順次送信しても良い。

詳細には、統計解析用コンピュータ S T は、要求者に関する 1 又は複数の「多型番地」及び「多型パターン」の組合せを受け取るステップと、受け取った「多型番地」及び「多型パターン」の組合せと一致する「多型番地」及び「多型パターン」の組合せを有する個体に関するゲノム関連情報をゲノム関連情報 D B 4 0 から検索するステップとを、検索の結果として 1 の個人（個体）に関するゲノム関連情報を要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報として特定するまで繰り返すことで、ゲノム関連情報 D B 4 0 に含まれる所定の個人（個体）に関するゲノム関連情報を要求者（要求者が送信した「個体関連情報」に関連する個人（個体））に関するゲノム関連情報として特定することができる。

一方、統計解析用コンピュータ S T は、複数の個体に関する複数の個体関連情報を、それぞれゲノム関連情報 D B 4 0 に含まれるゲノム関連情報と関連付けて格納した個体関連情報 D B 4 1 を用いて、当該複数の個体関連情報と当該複数のゲノム関連情報とを統計的に処理して、塩基配列関連情報を意味づける意味情報及び/又は当該意味情報に関連する情報を創出することができる。

ここで統計的処理としては、遺伝統計学的手法として知られている手法を適用した処理を意味し、従来知られている各種プログラム及びアルゴリズムを適用して行うことができる。

以下に一例を示す。まず、ゲノム関連情報 D B 4 0 のその時点における全データ（全「整理No.」に関連付けられた全データ）から、多型番地毎に当該多型番地においてとりうる多型パターンのそれぞれの出現頻度を算出する。算出結果は、例えば、

図 1 1 に示すように、行方向に多型番地を並べ、列方向に多型パターン毎の出現頻度を並べて示されるマトリックスとして表される。なお、図 1 1 では、多型番地「000001」における多型パターン「多型 1」の出現頻度は100人中50人であることを表している。

- 5 次に、ゲノム関連情報DB 4 0 と個体関連情報DB 4 1 とを用いて、所定の個体関連情報に関して、多型番地毎に当該多型番地においてとりうる多型パターンのそれぞれの出現頻度を算出する。

- 具体的には、個体関連情報DB 4 1 から、例えば、「嗜好a（例：赤色よりも青色の方が好きだ）」について該当する（結果欄に「○」が付いている）「整理No.」だけを抽出した上で、ゲノム関連情報DB 4 0 から前記抽出した「整理No.」を検索し、検索したデータのなかから、多型番地毎に当該多型番地においてとりうる多型パターンのそれぞれの出現頻度を算出する。算出結果は、例えば、図 1 2 に示すように、個体関連情報毎（「嗜好」毎）に行方向に多型番地を並べ、列方向に多型パターン毎の出現頻度を並べて示されるマトリックスとして表される。なお、図 1 2 では、「嗜好a」について該当する場合に、多型番地「000001」における多型パターン「多型 1」の出現頻度は50人中45人であることを表している。

- 次に、図 1 1 に示した出現頻度及び図 1 2 に示した出現頻度の結果を用いて、所定の個体関連情報に該当するか否かに関わらず算出した、多型番地毎の当該多型番地においてとりうる多型パターンのそれぞれの出現頻度（図 1 1）と、所定の個体関連情報に該当する場合に限って算出した、多型番地毎の当該多型番地においてとりうる多型パターンのそれぞれの出現頻度（図 1 2）とを、互いに比較し相対的な値を求める。

- 具体的には、例えば、図 1 2 に示した出現頻度の結果として、「嗜好a」について該当する場合における、多型番地「000001」における多型パターン「多型 1」の出現頻度は50人中45人と表される当該45人という値を、図 1 1 に示した出現頻度の結果を表す、ゲノム関連情報DB 4 0 のその時点における全データ（全「整理No.」に関連付けられた全データ）における多型番地「000001」における多型パターン「多型 1」の出現頻度100人中50人の当該50人で除した割合値（90%）を、前記相対的な値として求める。同様に、多型番地「000001」における多型パターン「多型 2」の場合の割合

値（12%）、多型番地「000001」における多型パターン「多型3」の場合の割合値（8%）を求める。その結果は、例えば、図13に示すように、個体関連情報毎に行方向に多型番地を並べ、列方向に多型パターン毎の割合値を並べて示されるマトリックスとして表される。

- 5 このときに、もし「嗜好a」が多型番地「000001」における多型パターンの相違による影響を全く受けないのであれば、上記算出した割合値（相対的な値）は、「多型1」、「多型2」、「多型3」のいずれにおいても同程度の値を示すはずである。この傾向は、図11及び図12に示した出現頻度を算出する際の母数が多いほど顕著に表れることが統計学的に知られている。
- 10 なお、図13に示した結果を求める際には、図11に示した出現頻度を用いる代わりに、例えば、統計解析用コンピュータST以外の外部機関が算出（統計的に処理）した、多型番地毎の当該多型番地においてとりうる多型パターンのそれぞれの出現頻度を示すデータ又は当該出現頻度を算出するための元データを用いても良い。
- 15 関連する個体の中に、個体関連情報DB41に含まれる個体関連情報に関連する個体が含まれていても含まれていなくても良い。

次に、図13に示した割合値を用いて、各多型番地毎に、当該多型番地において取りうる各多型パターンの相対値を算出する。具体的には、図13に示した結果では、例えば、「嗜好a」において、多型番地「000001」の多型パターン「多型1」における割合値は90%、多型パターン「多型2」における割合値は12%、多型パターン「多型3」における割合値は8%と表されるが、そのうち最小の割合値を示す多型パターン（「多型3」）の割合値（8%）を基準にして、各多型パターンにおける割合値を「多型3」の割合値（8%）で除す。すなわち多型パターン「多型1」の相対値は $11.25 (=90/8)$ 、「多型2」の相対値は $1.5 (=12/8)$ 、「多型3」の相対値は $1.0 (=8/8)$ となり、最小の割合値を示す「多型3」の割合値を基準にした倍率として、各多型パターンの相対値を求める。その結果として、例えば図14に示すように、個体関連情報毎に行方向に多型番地を並べ、列方向に多型パターン毎の相対値を並べて示されるマトリックスとして表される。

20

25

このとき、もし「嗜好a」が多型番地「000001」における多型パターンの相違による

影響を全く受けないのであれば、上記算出した倍率（相対値）は、「多型 1」、「多型 2」、「多型 3」のいずれにおいても 1.0 に近い値を示すはずである。この傾向は、図 1 1 及び図 1 2 に示した出現頻度を算出する際の母数が大きいほど顕著に表れることが統計学的に知られている。

- 5 次に、図 1 4 に示した結果に基づいて、相対値が所定の値を超えた場合に、当該相対値を示す多型番地における多型パターンの相違が、所定の個体関連情報に（該当するか否かに）関係していると判断する。具体的には、例えば、図 1 4 に示した結果において、倍率（相対値）が閾値を超えた場合に、当該倍率を示す多型番地とその倍率とを抽出し、出力する。なお、当該閾値は、例えば、図 1 4 に示した、所定の
10 の個体関連情報に該当する場合の全ての相対値を概観したうえで、統計的に算出することができる。

なお、図 1 4 に示した相対値の算出方法によっては、相対値が所定の値を下回った場合に当該相対値の多型番地における多型パターンの相違が、所定の個体関連情報に（該当するか否かに）関係していると判断してもよい。

- 15 また、抽出した倍率（相対値）が相対的に大きいほど、当該抽出した倍率を示す多型番地における多型パターンの相違が、所定の個体関連情報に（該当するか否かに）より強く関係していると推定できる。すなわち、所定の個体関連情報に（該当するか否かに）先天的（遺伝的）要素が強いと推定される。

- さらにまた、図 1 1 及び図 1 2 に示した出現頻度から図 1 4 に示した結果を得る
20 までの処理を、複数の個体関連情報について同時に並行して行った場合には、それぞれの個体関連情報について抽出された、それぞれの多型番地とその倍率（相対値）との傾向から、どの個体関連情報とどの個体関連情報とが互いに関連性があるのかが推定できる。

- 以上のように、図 1 1 及び図 1 2 に示した出現頻度から図 1 4 に示した結果を得て、
25 当該結果から判断を導くまでの処理を行うことにより、今まで先天的（遺伝的）影響との関連性が希薄と思われていた個体関連情報についても遺伝的影響を受けているか否かを調べることができる。すなわち、各人（個体）より入手する様々な種類の個体関連情報を用いることにより、従来、先天的（遺伝的）影響との関連性が希薄と思われた個体の性質等と遺伝との関連性を推定することができる。

また、図 1 1 及び図 1 2 に示した出現頻度から図 1 4 に示した結果を得て、当該結果から判断を導くまでの処理は、所定の個体関連情報（属性）に該当するか否かに拘わらず算出した多型パターン毎の出現頻度を母数（図 1 1）にしているので、個体関連情報（属性）がどれだけ遺伝の影響を受けているかを定量的に表すことに向いている。これに対して、通常行われる相関解析の方法では、個体関連情報（属性）に該当する集団と該当しない集団とを比較しているので、個体関連情報（属性）が人の性格のように遺伝による影響が比較的少ないと思われるものの場合に遺伝的影響の度合いを表すのが困難となる。すなわち、人の性格のような個体関連情報（属性）の場合、個々人の“かくありたい”という願望などが起因して当該個体
5 関連情報（属性）に該当しているか否かを判断すること自体が難しいため、母数に誤りが含まれることがあり正確に比較することが難しくなるからである。

なお、図 1 4 に示した結果において、倍率（相対値）が閾値を超えた場合に当該倍率を示す多型番地とその倍率とを抽出する際に、図 1 1、図 1 2 及び図 1 3 に示した出現頻度や割合値及びそれらの算出根拠となった元の数値も抽出して併せて出力することが望ましい。この場合、図 1 1、図 1 2 及び図 1 3 に示した出現頻度や割合値及びそれらの算出根拠となった元の数値を、図 1 4 に示した結果を使用する側で、データの信憑性を判断する材料として使用することができる。

その後、統計解析用コンピュータ S T は、図 1 4 に示した結果から抽出した多型番地とその倍率及び/又は併せて抽出した算出根拠となる元の数値を用いることにより、所定の個体関連情報と所定の多型番地との相関関係及び多型パターンの種類による所定の個体関連情報への影響の差異という知見を得ることができる。そして統計解析用コンピュータ S T は、得た知見を基に意味情報及び/又は当該意味情報に関連する情報を創出することができる。例えば、図 1 4 に示した結果から抽出した「倍率」は、そのまま意味情報として使用することが可能である。

20 5 なお、統計解析用コンピュータ S T は、図 1 4 に示した倍率（相対値）の算出結果及び/又は図 1 4 に示した算出結果から導かれた知見に基づいて創出した意味情報及び/又は当該意味情報に関連する情報を、逐次記録したデータベース（以下、参照用 D B と称す）を構築することができる。

また、所定の個体関連情報（性質）に関して所定の多型番地との相関関係が見い

だせない場合であっても、以下のようにして、1又は複数の他の個体関連情報（性質）との組合せと所定の多型番地との相関関係を見いだすことができる。

例えば、「多汗症である」という個体関連情報（性質）と相関関係がある多型番地が見いだせなかったとする。また、「多汗症である」性質に関連性がある性質、

- 5 例えば「よく水分を摂取する」性質及び「塩辛いものを好む」性質については、それぞれ「多型番地000001」及び「多型番地000101」に相関関係が見いだされたとする。ここで、所定の性質と他の性質との関連性とは、所定の性質を示す個体が他の性質を示す確率が所定の値、例えば80%以上であることを意味する。前記関連性は、
10 例えば、所定の性質と他の性質を含むアンケートを実施した集計結果から導き出すことができる。例えば、所定の性質を有すると答えた回答者の中で、他の性質を有すると答えた回答者が80%以上の場合は、所定の性質と他の性質との間に関連性があるとみなすことができる。

- さらに、「よく水分を摂取する」性質を示す個体のうち80%の個体は、「多型番地000001」において「多型1」を有しており、「塩辛いものを好む」性質を示す個
15 体のうち70%の個体は、「多型番地000101」において「多型2」を有していたとする。

- 以上の前提条件の基では、「多汗症である」という性質について、「多型番地000001」及び「多型番地000101」との間に間接的な相関関係を見いだすことができる。すなわち、「多型番地000001」において「多型1」を有し、「多型番地
20 000101」において「多型2」を有する個体のうち、56% ($0.8 \times 0.7 = 0.56$) は「多汗症である」性質を示すことが間接的に見いだせる。

- 以上によって、「多型番地000001」において「多型1」を有し、「多型番地000101」において「多型2」を有する場合に、「多汗症である場合が多い（例えば「指数56」）」という意味情報を創出することができる。さらに、創出した意味
25 情報から、例えば、「生活留意情報」等の当該意味情報に関連する情報を導き出すこともできる。

統計解析用コンピュータSTは、以上のようにして、創出した意味情報及び/又は当該意味情報に関連する情報を共用コンピュータ2に提供することができる。共用コンピュータ2は、統計解析用コンピュータSTから提供された意味情報に基づい

てメインDB 1 4を構築したり、統計解析用コンピュータ S Tから提供された当該意味情報に関連する情報に基づいて情報提供用データベースを構築することができる。

また、統計解析用コンピュータ S Tは、上述したように創出した意味情報及び/又は当該意味情報に関連する情報等を逐次記録して「参照用 D B」を構築し、当該

5 「参照用 D B」をメインDB 1 4として使用することによって、自らが共用コンピュータ 2として動作することもできる。

共用コンピュータ 2は、構築したメインDB 1 4を用いて、以下のようにして利用者に対して所定の疾病の罹患可能性等の意味情報を提供することができる。すなわ

ち、共用コンピュータ 2は、利用者が「物品及び/又はサービスの要求」として、例
10 えば、所定の疾病に関する自分の罹患可能性を教えて欲しいと要求する場合に、当該罹患可能性等の意味情報を提供できる。

なお、本情報処理システムにおいて、「物品及び/又はサービス」としては、所定の疾病の罹患可能性に限定されず、例えば、個人（個体）の体質に適合した医薬品、食品及び嗜好品等の物品や、個人（個体）の体質・性質に適合した情報等のサ
15 ービスを含む意味である。

共用コンピュータ 2が利用者に対して所定の疾病の罹患可能性を提供する情報処理システムにおいては、共用コンピュータ 2のメモリー 7に記録された処理プログラム 1 3及び個人用コンピュータ 3のメモリー 2 3に記録された処理プログラム 2
7が例えば、図 1 5及び図 1 6に示すようなフローチャートに従って情報処理動作
20 する。なお、図 1 5及び図 1 6に示すフローチャートにおいて、「（共）」と記載したステップは共用コンピュータ 2における処理を意味し、「（個）」と記載したステップは個人用コンピュータ 3における処理を意味している。

本情報処理システムは、ゲノム関連情報記録媒体 2 4を所持する各個人が個人用コンピュータ 3を用いて通信回線網 1を介して共用コンピュータ 2にアクセスし、
25 共用コンピュータ 2のメインDB 1 4に記録されている意味情報を利用するシステムである。なお、本情報処理システムは、複数人のゲノム関連情報 2 8がそれぞれ記録されたゲノム関連情報記録媒体 2 4を用い、各個人がゲノム関連情報記録媒体 2 4にアクセスするようなシステムであつてもよい。

この場合、まず、ステップ A 1（SA1）で、要求者が本システムを利用するにあた

り、メモリー 23 に記録されている処理プログラム 27 を起動する。処理プログラム 27 によって、個人用コンピュータ 3 の読取り装置 25 を駆動してゲノム関連情報記録媒体 24 にアクセスし、ゲノム関連情報記録媒体 24 においてデータ I として記録されている「Gno.」を読み出す。読み出した「Gno.」は、メモリー部 26 に格納する。

次に、ステップ A2 (SA2) では、処理プログラム 27 によって表示装置 22 に表示された画面イメージに基づいて、要求者が提供を受けたい情報、例えば、「大腸がんの罹患可能性」(要求情報)を個人用コンピュータ 3 に入力するとともに、個人用コンピュータ 3 から通信回線網 1 を経由して共用コンピュータ 2 に「大腸がんの罹患可能性」及び「Gno.」を送信する。或いは、個人用コンピュータ 3 から通信回線網 1 を経由して共用コンピュータ 2 に対して、「大腸がんの罹患可能性」及び「Gno.」を書き込む。

次に、ステップ A3 (SA3) では、共用コンピュータ 2 が「大腸がんの罹患可能性」及び「Gno.」を受信する。受信した「大腸がんの罹患可能性」及び「Gno.」は、メモリー部 A10 に要求情報として格納する。

次に、ステップ A4 (SA4) では、要求情報を受信すると、メモリー 7 に記録されている処理プログラム 13 を起動してメイン DB14 にアクセスする。なお、この処理プログラム 13 は、共用コンピュータ 2 における処理を行うものである。

次に、ステップ A5 (SA5) では、処理プログラム 13 に従って、メイン DB14 に記録されている「分類(疾患名)」を検索し、要求された「大腸がんの罹患可能性」(大腸がん)と一致するものを抽出する。

ステップ A6 (SA6) では、メイン DB14 に記録されているデータのなかから「大腸がんの罹患可能性」と一致した「分類(疾患名)」(大腸がん)に関連づけられた「多型番地」を読み出す。読み出した「多型番地」は、メモリー部 A10 に要求情報に関連づけた位置情報として格納する。すなわち、メモリー部 A10 には、所定の「Gno.」に対して「大腸がんの罹患可能性」及び「多型番地」が記録されることとなる。

次に、ステップ A7 (SA7) では、メモリー部 A10 に記録されている「Gno.」及び「多型番地」を個人用コンピュータ 3 に送信するとともに、送信する「多型番

地」に対応する「多型パターン」を提出する命令情報を個人用コンピュータ 3 に送信する。また、このとき、要求情報の種類によっては、必要に応じて既往症や特徴等の付加的な情報の提出を命令してもよい。

次に、ステップA 8 (SA8) では、共用コンピュータ 2 から送信された「Gno.」、
5 「多型番地」及び命令情報を受信する。受信した「Gno.」及び「多型番地」は、メモリー部 2 6 に記録される。

次に、ステップA 9 (SA9) では、受信した命令情報に従って、ゲノム関連情報記録媒体 2 4 に記録されているデータIIにアクセスする。ステップA 1 0 (SA10) では、
10 処理プログラム 2 7 に従ってゲノム関連情報記録媒体 2 4 に記録されているデータIIを検索し、命令された多型番地の多型パターンを読み出し、多型番地と多型パターンとを関連づけてメモリー部 2 6 に記録する。このとき、データIに対してアクセスし、ステップA 8 で受信した「Gno.」が正しいか否かを確認することが好ましい。
また、ステップA 1 0 では、多型パターンのほかにデータIII、データIV及びデータV
15 に記録されている付加的な情報も同時に読み出し、必要に応じてメモリー部 2 6 に記録してもよい。

次に、ステップA 1 1 (SA11) では、メモリー部 2 6 に一時的に記録した多型番地に関連付けられた多型パターン及び必要に応じて記録された付加的な情報を、
「Gno.」とともに通信回線網 1 を介して共用コンピュータ 2 に対して出力する。ステップA 1 2 (SA12) では、多型番地に関連付けられた多型パターン及び必要に応じて
20 て記録された付加的な情報を共用コンピュータ 2 で受信し、受信した多型パターンを多型番地と関連付けてメモリー部 A 1 0 に記録する。

また、本例では、ステップA 7 において、共用コンピュータ 2 が「多型パターン」の提出を命令する命令情報を送出し、ステップA 1 0 において、個人用コンピュータ 3 は命令情報に従って多型パターンをゲノム関連情報記録媒体 2 4 から読み出して
25 いる。しかしながら、本システムは、ステップA 7 において当該命令情報を送出しないシステムであってもよい。この場合、ステップA 1 0 において、個人用コンピュータ 3 は、処理プログラム 2 7 に従って、ステップA 8 で受信した多型番地に基づいてデータIIを検索し、受信した多型番地の多型パターンを読み出す。そして、個人用コンピュータ 3 は、ステップA 1 1 で多型パターン等を共用コンピュータ 2 に対して

出力する。この場合でも、共用コンピュータ 2 は、ステップ A 1 2 において、「大腸がんの罹患可能性」と一致した「分類（疾患名）」に関連づけられた「多型番地」の多型パターンを得ることができる。

次に、ステップ A 1 3（SA13）では、メイン DB 1 4 にアクセスし、受信した多型番地及び多型パターンと一致するものを検索する。具体的には、メイン DB 1 4 において、一つの多型番地に対して複数の多型パターンが記録されており、受信した多型番地及びその多型パターンがメイン DB 1 4 においてどの多型パターンに一致しているのかを検索する。

次に、ステップ A 1 4（SA14）では、処理プログラム 1 3 に従って、受信した多型パターンと一致した多型パターンに関連づけられている大腸がんに対する罹患可能性を読み出す。すなわち、ステップ A 1 4 では、要求者が提出した多型番地及び多型パターンに従って、要求者の大腸がんに対する罹患可能性を読み出すことができる。読み出した罹患可能性は、要求者の「Gno.」と関連づけてメモリー部 A 1 0 に格納する。このとき、大腸がんに対する罹患可能性を、付加的な情報により補正したかたちで格納してもよいし、付加的な情報から得られるその他の情報を要求者の「Gno.」に関連づけて格納しても良い。

次に、ステップ A 1 5（SA15）では、メモリー部 A 1 0 に格納した要求者の「Gno.」及び罹患可能性を意味情報として、通信回線網 1 を介して個人用コンピュータ 3 に対して送信する。ステップ A 1 6（SA16）では、個人用コンピュータ 3 が要求者の「Gno.」及び罹患可能性（意味情報）を受信する。受信した意味情報は、メモリー部 2 6 に記録される。

次に、ステップ A 1 7（SA17）では、処理プログラム 2 7 に従って、メモリー部 2 6 に記録された意味情報から大腸がんに対する罹患可能性を表示装置 2 2 に表示する。なお、ステップ A 1 5 からステップ A 1 7 の代わりに共用コンピュータ 2 が処理プログラム 1 3 に従って意味情報を表示する画面を読み出し（作成し）、通信回線網 1 を経由して個人用コンピュータ 3 の表示装置 2 2 に表示させることもできる。この場合においても、共用コンピュータ 2 から個人用コンピュータ 3 に対して意味情報が送信されたものとする。これにより、要求者は、ゲノム関連情報記録媒体 2 4 に記録したゲノム関連情報 2 8 を用いて大腸がんに対する罹患可能性を得ることが

できる。

以上のように、本システムにおいては、個人の多型パターンを多型番地と関連づけて記録したゲノム関連情報記録媒体 2 4 を用いることによって、メインDB14に記録された意味情報を多型番地を介在させて個人が利用することができる。言い換えれば、本システムを利用する個人は、意味情報をゲノム関連情報記録媒体に記録しておく必要はなく、多型番地と多型パターンとを関連づけたゲノム関連情報 2 8 を所有するだけで、様々な意味情報を得ることができる。

特に、意味情報は、上述したように、その種類が増加するとともに訂正が行われるため、メインDB14を更新することによってより精度が高く、且つ、幅広い情報を含むものとなる。本システムによれば、このような意味情報の増加及び訂正等に従ってメインDB14を更新することによって、個人が最新の意味情報を利用することができる。

さらに、ゲノム関連情報 2 8 を記録したゲノム関連情報記録媒体 2 4 を用いることによって、利用者は本システムを利用するたび毎にゲノム関連情報を得るための検査をする必要がない。すなわち、利用者は、一旦、ゲノム関連情報記録媒体 2 4 を作製すれば、以降は本システムを利用して最新の意味情報を得ることができる。

ゲノム関連情報 2 8 を記録したゲノム関連情報記録媒体 2 4 を利用者自身が保有する場合、本人のゲノム関連情報 2 8 を外部の機関に委託して保管させる際の不安や当該機関に対する不正アクセスによりゲノム関連情報 2 8 が流出するといった危険性を回避することができる。一方、ゲノム関連情報記録媒体 2 4 に複数の個人に関する複数のゲノム関連情報 2 8 を記録して外部の機関に委託して保管させる場合、個々人がゲノム関連情報記録媒体 2 4 を保有する場合と比較して、ゲノム関連情報記録媒体 2 4 の取り扱いの不手際やゲノム関連情報記録媒体 2 4 の損失といった事態を防止することができる。

特に、この図 1 5 及び図 1 6 に示したフローチャートに従えば、ゲノム関連情報記録媒体 2 4 に記録したゲノム関連情報 2 8 の全てを通信回線網 1 を介して出力する必要がなく、提出命令を受けた一部のゲノム関連情報 2 8 のみを出力すればよい。したがって、本システムによれば、機密性の高い個人特有の多型番地及び多型パターンの漏洩を防止することができる。

また、この図 1 5 及び図 1 6 に示したフローチャートに従えば、共用コンピュータ 2 において要求者に提供する意味情報を得ているため、個人用コンピュータ 3 においてメインDB 1 4 に記録されている情報を取り扱う必要がない。したがって、この図 1 5 及び図 1 6 に示したフローチャートに従えば、個人用コンピュータ 3 の情報処理能力が比較的低くても、十分に所望の意味情報を得ることができる。さらに、個人用コンピュータ 3 においてメインDB 1 4 に記録されている情報を取り扱う必要がないため、個人用コンピュータ 3 の処理プログラム 2 7 を、ゲノム関連情報記録媒体 2 4 を装着するカードドライブ等に併せて規格化しやすくなる。

ところで、本情報処理システムにおいては、共用コンピュータ 2 のメモリー 7 に記録された処理プログラム 1 3 及び個人用コンピュータ 3 のメモリー 2 3 に記録された処理プログラム 2 7 が例えば、図 1 7 に示すようなフローチャートに従って情報処理動作するものであってもよい。なお、図 1 7 に示すフローチャートにおいても、「(共)」と記載したステップは共用コンピュータ 2 における処理を意味し、「(個)」と記載したステップは個人用コンピュータ 3 における処理を意味している。

ここでは、まず、ステップB 1 (SB1) では、要求者が本システムを利用するにあたり、メモリー 2 3 に記録されている処理プログラム 2 7 を起動する。処理プログラム 2 7 によって、個人用コンピュータ 3 の読取り装置 2 5 を駆動してゲノム関連情報記録媒体 2 4 にアクセスし、ゲノム関連情報記録媒体 2 4 においてデータIとして記録されている「Gno.」を読み出す。読み出した「Gno.」は、メモリー部 2 6 に格納する。

次に、ステップB 2 (SB2) では、処理プログラム 2 7 によって表示装置 2 2 に表示された画面イメージに基づいて、要求者が提供を受けたい情報、例えば、「大腸がんの罹患可能性」(要求情報)を個人用コンピュータ 3 に入力するとともに、個人用コンピュータ 3 から通信回線網 1 を経由して共用コンピュータ 2 に「大腸がんの罹患可能性」及び「Gno.」を送信するとともに、メインDB14の「分類(疾患名)」が大腸がんである「多型番地」と当該「多型番地」に関連付けられた全ての「多型パターン」と当該全ての「多型パターン」それぞれを意味づける「罹患可能性」との提出を要求する。すなわち、要求者は、ステップB2において、メインDB14

の「分類（疾患名）」が大腸がんである「多型番地」と当該「多型番地」に関連付けられた全ての「多型パターン」と当該全ての「多型パターン」それぞれを意味づける「罹患可能性」とからなる情報を要求する。

次に、ステップB3（SB3）では、共用コンピュータ 2 が上記要求情報を受信する。

- 5 共用コンピュータ 2 は、要求情報を受信すると処理プログラム 1 3 を起動する。そして、ステップB4（SB4）で、処理プログラム 1 3 に従ってメインDB14にアクセスする。

次に、ステップB5（SB5）では、処理プログラム 1 3 に従って、メインDB14に記録されている「分類（疾患名）」を検索し、要求された「大腸がんの罹患可能性」

- 10 （大腸がん）と一致するものを抽出する。ステップB6（SB6）では、処理プログラム 1 3 に従って、メインDB14にアクセスし、「大腸がんの罹患可能性」と一致する「分類（疾患名）」（大腸がん）に関連づけられた「多型番地」、当該多型番地に関連づけられた全ての「多型パターン」及び全ての多型パターンにおける「罹患可能性」を読み出す。読み出した「多型番地」、「多型パターン」及び「罹患可能性」は、メモリー部 A 1 0 に要求情報に関連づけて格納する。すなわち、メモリー部 A 1 0 には、所定の「Gno.」に対して「多型番地」、「多型パターン」及び「罹患可能性」が記録されることとなる。

- 次に、ステップB7（SB7）では、メモリー部 A 1 0 に記録されている「Gno.」、「多型番地」、「多型パターン」及び「罹患可能性」を、通信回線網 1 を介して個人用コンピュータ 3 に対して送信する。ステップB8（SB8）では、共用コンピュータ 2 から送信された「Gno.」、「多型番地」、「多型パターン」及び「罹患可能性」を受信する。受信した「Gno.」、「多型番地」、「多型パターン」及び「罹患可能性」は、メモリー部 2 6 に記録される。

- 次にステップB9（SB9）では、処理プログラム 2 7 に従い、ゲノム関連情報記録媒体 2 4 に記録されているデータIIにアクセスする。このとき、ゲノム関連情報記録媒体 2 4 に記録されているデータIにもアクセスし、受信した「Gno.」が正しいか否かを確認することが好ましい。

次に、ステップB10（SB10）では、処理プログラム 2 7 に従って、ゲノム関連情報 2 8 から、受信した「多型番地」と一致する多型番地における多型パターンを抽出

する。そして、ステップB10 (SB10) では、受信した多型番地に関連づけられた全ての「多型パターン」のうちで、抽出した多型パターンと一致するものを検索する。

ステップB11 (SB11) では、受信した多型番地に関連づけられた全ての「多型パターン」のうちで一致した多型パターンに関連づけられた「罹患可能性」を抽出するとともに、抽出した「罹患可能性」を出力する。これにより、要求者は、大腸がんに対する罹患可能性（意味情報）を得ることができる。このとき、ステップB11では、データIII、データIV及びデータVに記録されている付加的な情報も同時に読み出し、大腸がんに対する罹患可能性を、付加的な情報により補正したかたちで出力してもよい。

特に、図17に示したフローチャートに従えば、ゲノム関連情報記録媒体24に記録したゲノム関連情報28を個人用コンピュータ3以外の外部に対して全く出力することがない。すなわち、ゲノム関連情報28は、ゲノム関連情報記録媒体24と個人用コンピュータ3との間でのみ、やり取りされる。したがって、本システムによれば、機密性の高い個人特有のゲノム関連情報28の漏洩をより確実に防止することができる。

ところで、本情報処理システムにおいては、共用コンピュータ2のメモリー7に記録された処理プログラム13及び個人用コンピュータ3のメモリー23に記録された処理プログラム27が例えば、図18に示すようなフローチャートに従って情報処理動作するものであってもよい。なお、図18に示すフローチャートにおいても、「(共)」と記載したステップは共用コンピュータ2における処理を意味し、「(個)」と記載したステップは個人用コンピュータ3における処理を意味している。

ここでは、まず、ステップC1 (SC1) で、要求者が本システムを利用するにあたり、メモリー23に記録されている処理プログラム27を起動する。処理プログラム27によって、個人用コンピュータ3の読取り装置25を駆動してゲノム関連情報記録媒体24にアクセスし、ゲノム関連情報記録媒体24においてデータIとして記録されている「Gno.」、データIIとして記録されている全ての「多型番地」及び「多型パターン」を読み出す。読み出した「Gno.」、「多型番地」及び「多型パターン」は、メモリー部26に格納する。

次に、ステップC2 (SC2) では、処理プログラム27によって表示装置22に表示された画面イメージに基づいて、要求者が提供を受けたい情報、例えば、「大腸がんの罹患可能性」(要求情報)を個人用コンピュータ3に入力するとともに、個人用コンピュータ3から通信回線網1を経由して共用コンピュータ2に「大腸がんの罹患可能性」と、メモリー部26に記録されている「Gno.」、「多型番地」及び「多型パターン」とを送信する。

次に、ステップC3 (SC3) では、共用コンピュータ2が「大腸がんの罹患可能性」、「Gno.」、「多型番地」及び「多型パターン」を受信する。受信した「大腸がんの罹患可能性」は要求情報としてメモリー部A10に記録され、「Gno.」、「多型番地」及び「多型パターン」も、メモリー部A10に格納される。共用コンピュータ2は、要求情報を受信すると処理プログラム13を起動する。そして、ステップC4 (SC4) では、処理プログラム13に従って、メインDB14にアクセスする。

次に、ステップC5 (SC5) では、処理プログラム13に従って、メインDB14に記録されている「分類(疾患名)」を検索し、要求された「大腸がんの罹患可能性」(大腸がん)と一致するものを抽出する。

ステップC6 (SC6) では、処理プログラム13に従って、メインDB14にアクセスし、メインDB14から「大腸がん」に分類された「多型番地」、当該多型番地に対する全ての「多型パターン」、及び当該多型パターンに対する「罹患可能性」を読み出す。読み出した「多型番地」、「多型パターン」及び「罹患可能性」は、メモリー部A10に格納される。

次に、ステップC7 (SC7) では、ステップC3で受信した「多型番地」及び「多型パターン」に基づいて、ステップC6でメモリー部A10に格納したデータを検索し、受信した「多型パターン」と一致した多型パターンに関連付けられた罹患可能性をメモリー部A10から抽出する。

ステップC8 (SC8) では、ステップC7の結果、すなわち、ステップC3で受信した情報に含まれる多型パターンがメインDB14のいずれの多型パターンと一致するかに基づいて抽出した罹患可能性を、通信回線網1を介して個人用コンピュータ3に対して送信する。このとき、共用コンピュータ2は、抽出した罹患可能性を要求者

の「Gno.」とともに送信する。

次に、ステップC9 (SC9) では、共用コンピュータ 2 から送信された「Gno.」及び「罹患可能性」(意味情報)を受信する。受信した「Gno.」及び「罹患可能性」は、メモリー部 2 6 に記録される。このとき、ゲノム関連情報記録媒体 2 4 に記録されているデータ I にアクセスし、受信した「Gno.」が正しいか否かを確認することができる。

次に、ステップC10 (SC10) では、処理プログラム 2 7 に従って、メモリー部 2 6 に記録された意味情報から大腸がんに対する罹患可能性を表示装置 2 2 に表示する。なお、ステップC8 からステップC10 の代わりに、共用コンピュータ 2 が処理プログラム 1 3 に従って意味情報を表示する画面を読み出し(作成し)、通信回線網1を経由して個人用コンピュータ 3 の表示装置 2 2 に表示させることもできる。この場合においても、共用コンピュータ 2 から個人用コンピュータ 3 に対して意味情報が送信されたものとする。これにより、要求者は、ゲノム関連情報記録媒体 2 4 に記録したゲノム関連情報 2 8 を用いて大腸がんに対する罹患可能性を得ることができる。

特に、図 1 8 に示したフローチャートに従えば、ゲノム関連情報記録媒体 2 4 に記録したゲノム関連情報 2 8 を全て共用コンピュータ 2 に対して出力し、共用コンピュータ 2 において要求者に提供する意味情報を得ている。このため、図 1 8 に示したフローチャートに従えば、個人用コンピュータ 3 と共用コンピュータ 2 との間での情報の授受が比較的少ない回数で、要求者が意味情報を得ることができる。したがって、この図 1 8 に示したフローチャートに従えば、個人用コンピュータ 3 の情報処理能力が比較的低くても、十分に所望の意味情報を得ることができるとともに、要求者にとっては非常に簡便に意味情報を得ることができる。

以上、説明したように本システムによれば、ゲノム関連情報記録媒体 2 4 及びメインDB 1 4 において、「多型番地」及びその「多型パターン」のみを規格化しておけば、それ以外の特別なデータの規格化を必要としないので、広範囲な産業に利用することができる。すなわち、物品或いはサービスを提供する側は、ゲノム関連情報記録媒体 2 4 を用いた情報提供に際して、多型パターンに対応する意味情報の規格化や、データ授受処理方法等の統一した規格を必要とせず、様々な方式で情報提

供することができる。

さらにまた、本システムによれば、メインDB 1 4をチェックすることで、第三者
或いは第三者機関は共用コンピュータ 2 に対する監視及び管理を容易に行うことが
できる。したがって、本システムは、意味情報を提供する側に対する例えば行政的
5 な管理を行うことができるため、意味情報を提供する側の健全性及び倫理管理を行
うことができる。

一方、本情報処理システムにおいては、ゲノム関連情報記録媒体からデータIIに
含まれる情報を除いたもの、すなわちデータI及び付加的にデータIII~Vのみを有す
る記憶媒体を用いても良い。この場合、データIIに含まれる情報は、通信回線網 1
10 を介して個人用コンピュータ 3 と接続された外部のデータベース（ゲノム関連情報
記録媒体）に記録しておく。このようなシステムの場合、例えば、上述したステッ
プA1 0において、通信回線網 1 を介して外部のデータベースにアクセスし、命令さ
れた多型番地の多型パターンを読み出し、多型番地と多型パターンとを関連づけて
メモリー部 2 6 に記録することができる。したがって、このようなシステムであつ
15 ても、図 1 5 及び図 1 6 に示したフローチャート、図 1 7 に示したフローチャート
及び図 1 8 に示したフローチャートと同様に、要求者は意味情報を得ることができ
る。

さらに、本情報処理システムにおいては、要求者がゲノム関連情報記録媒体 2 4
及び前記ゲノム関連情報記録媒体からデータIIに含まれる情報を除いた記録媒体の
20 いずれも有さず、通信回線網 1 を介して個人用コンピュータ 3 と接続したゲノム関
連情報記録媒体 2 4 を備えるものであっても良い。このようなシステムの場合、要
求者は、通信回線網 1 を介してゲノム関連情報記録媒体 2 4 にアクセスし、ゲノム
関連情報記録媒体 2 4 に記録された「多型番地」及び「多型パターン」等の情報を
個人用コンピュータ 3 にダウンロードできる。なお、この場合、ゲノム関連情報記
25 録媒体 2 4 は、複数の個人に関するゲノム関連情報を個人毎（「Gno.」毎）に記録
したものであっても良い。

さらにまた、本発明は、上述したような共用コンピュータ 2 がメインDB 1 4 を有
するような構成に限定されず、例えば、共用コンピュータ 2 と通信回線網 1 を介し
て接続されたメインDB 1 4 を備える情報処理システムにも適用される。この場合、

共用コンピュータ 2 は、図 1 5 及び図 1 6 に示したフローチャート、図 1 7 に示したフローチャート或いは図 1 8 に示したフローチャートにおいて、メインDB 1 4 に対して通信回線網 1 を介してアクセスする。この場合でも、本情報処理システムによれば、図 1 5 及び図 1 6 に示したフローチャート、図 1 7 に示したフローチャート或いは図 1 8 に示したフローチャートに従って要求者が所望の意味情報を得ることができる。

特に、この場合、共用コンピュータ 2 は、異なる機関又は団体が有する複数のメインDB 1 4 に対して通信回線網 1 を介してアクセスし、これら複数のメインDB 1 4 に含まれる意味情報を使用して、要求者に対する情報提供を行うことが可能となる。すなわち、本情報処理システムにおいては、図 1 5 及び図 1 6 に示したフローチャートにおけるステップA 5 で、図 1 7 に示したフローチャートにおけるステップB 5 で、或いは図 1 8 に示したフローチャートにおけるステップC 5 で、共用コンピュータ 2 が大腸がんの罹患可能性に関する情報を意味情報として有する様々なメインDB 1 4 にアクセスする。これにより、本情報処理システムによれば、要求者は、様々なメインDB 1 4 に含まれる情報に基づいて、大腸がんの罹患可能性に関する情報を得ることができる。

また、本システムは、図 1 5 及び図 1 6 に示したフローチャート、図 1 7 に示したフローチャート、或いは図 1 8 に示したフローチャートにおいて、共用コンピュータ 2 が、いわゆるエージェントに対して、少なくとも個人用コンピュータ 3 から受け取った要求情報を送信し、意味情報（本例においては、「大腸がんに関する罹患可能性」）を、当該エージェントを介して得るものであってもよい。

本明細書で引用した全ての刊行物、特許および特許出願をそのまま参考として本明細書にとり入れるものとする。

産業上の利用の可能性

以上、詳細に説明したように、本発明によれば、個体間における塩基配列情報の相違を有効に利用して各個体にとって有益な意味情報及び/又は当該意味情報に関連する情報を提供できる情報処理システムを構築することができ、特に、当該意味情報を創出することができる情報処理システムを構築することができる。

請求の範囲

1. 所定の個体に関する塩基配列関連情報を受け取るステップ a と、
塩基配列における位置を意味する位置情報及び当該位置情報に対応する塩基配列
5 関連情報が互いに関連付けられたセットが複数含まれた塩基配列関連情報群が個体
毎に格納された記憶装置から、前記受け取った塩基配列関連情報と一致性がある塩
基配列関連情報を含んだ塩基配列関連情報群を特定するステップ b と、
を有する塩基配列に関する情報処理方法。
2. 前記ステップ a では、予め決まった位置情報に対応する塩基配列関連情報を
10 受け取ることを特徴とする請求項 1 記載の情報処理方法。
3. 前記ステップ b で特定した塩基配列関連情報群が複数ある場合には、
所定の個体に関する塩基配列関連情報の受け取りと、複数の塩基配列関連情報群
の中で、当該受け取った塩基配列関連情報と一致性がある塩基配列関連情報を含ん
だ塩基配列関連情報群の特定とを、唯一の塩基配列関連情報群を特定するまで繰り
15 返すステップ c を更に有することを特徴とする請求項 1 記載の情報処理方法。
4. 前記所定の個体に関する個体関連情報を受け取るステップ d を更に有し、
前記ステップ a、ステップ b 及びステップ d を複数の個体に関して行い、当該複
数の個体に関する複数の塩基配列関連情報群と当該複数の個体に関する複数の個体
関連情報とをそれぞれ関連付けたデータベースを構築することを特徴とする請求項
20 1 記載の情報処理方法。
5. 前記データベースに含まれる複数の個体に関する複数の個体関連情報と、
前記データベースに含まれる若しくは含まれない複数の個体に関する複数の塩基配
列関連情報群又は当該複数の塩基配列関連情報群を統計的に処理した結果とを統計
的に処理して、塩基配列関連情報を意味づける意味情報及び/又は当該意味情報に関
25 連する情報を創出することを特徴とする請求項 4 記載の情報処理方法。
6. 複数の個体に関する複数の塩基配列関連情報群と当該複数の個体に関する
複数の個体関連情報とをそれぞれ関連付けて構築されたデータベースに含まれる複
数の個体に関する複数の個体関連情報と、前記データベースに含まれる若しくは含
まれない複数の個体に関する複数の塩基配列関連情報群又は当該複数の塩基配列関

連情報群を統計的に処理した結果とを統計的に処理して創出された塩基配列関連情報を意味づける意味情報及び/又は当該意味情報に関連する情報を取得し、

- 取得した意味情報及び/又は当該意味情報に関連する情報を用いて、物品及び/又はサービスの要求情報に応じた意味情報及び/又は当該意味情報に関連する情報を提供するための記憶装置の記憶内容を構築することを特徴とする塩基配列に関する情報処理方法。
- 5

図1

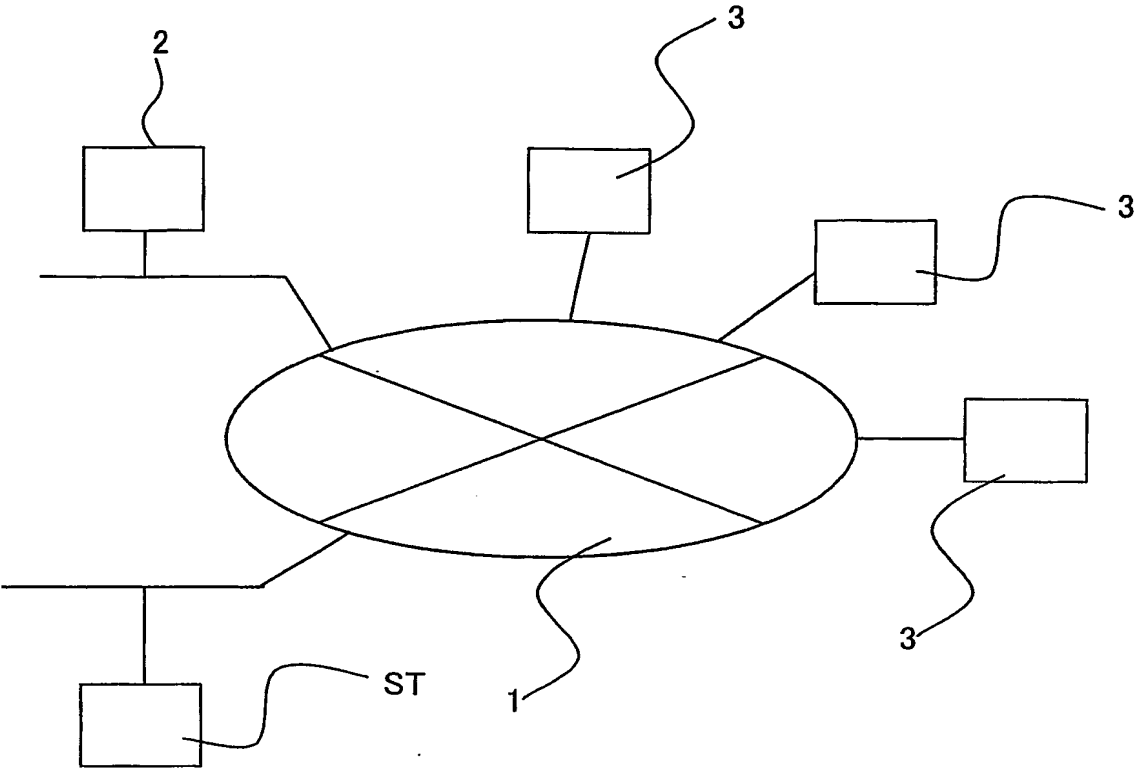


図2

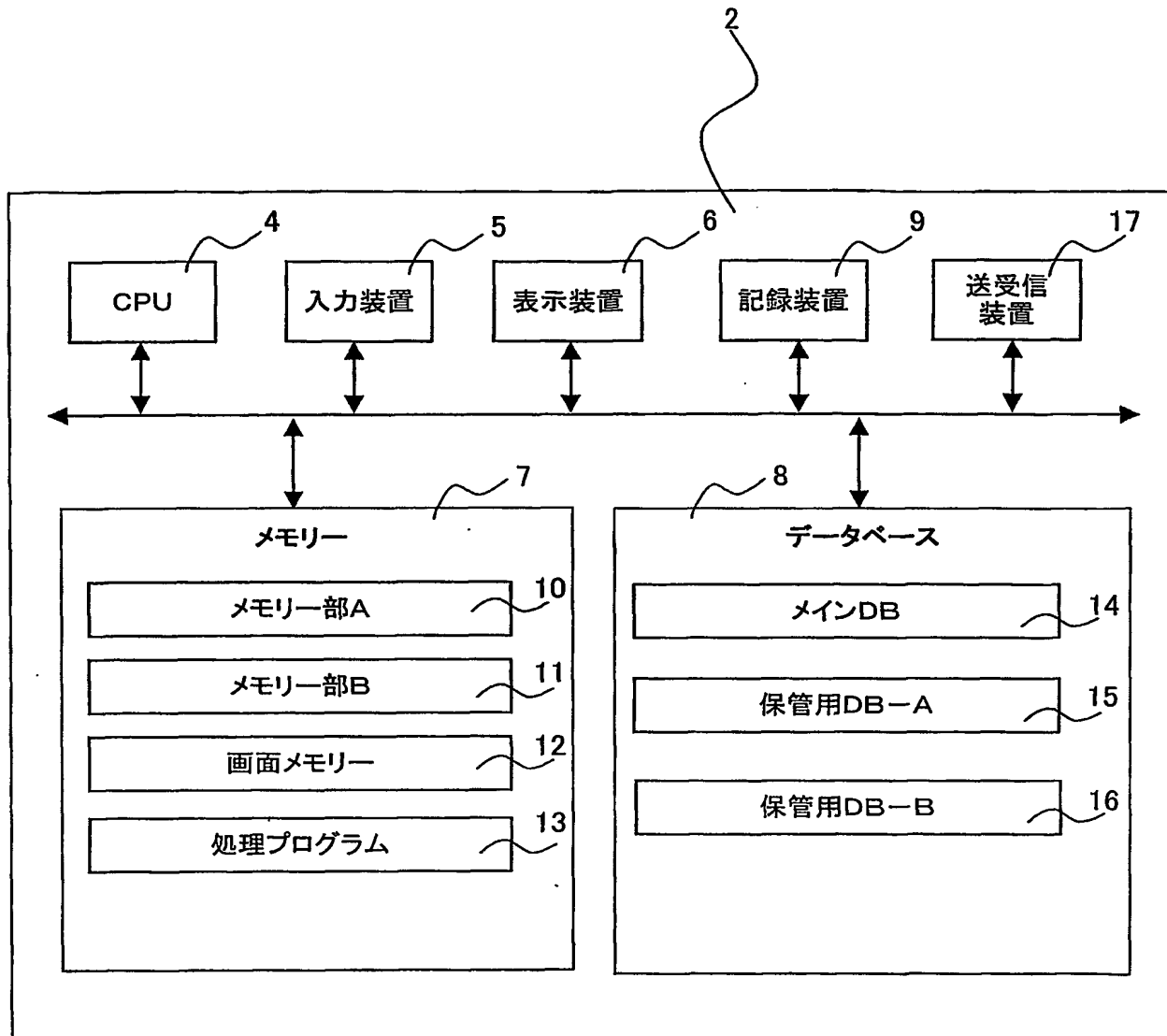


図3

多型番地	多型分類	多型パターン	分類 (疾患名)	多型パターンに対する 注釈情報 (罹患可能性)	公開レベル (公開可否)
123456	SNP	多型1	高血圧症	a	○
123456	SNP	多型2	高血圧症	b	○
123456	SNP	多型3	高血圧症	c	○
223456	SNP	多型1	大腸がん	イ	○
223456	SNP	多型2	大腸がん	ロ	○
223456	SNP	多型3	大腸がん	ハ	○
234567	SNP	多型1	胃がん	d	○
234567	SNP	多型2	胃がん	e	○
234567	SNP	多型3	胃がん	f	○
334567	SNP	多型1	喘息	二	○
334567	SNP	多型2	喘息	ホ	○
334567	SNP	多型3	喘息	ヘ	○
345678	SNP	多型1	糖尿病	g	○
345678	SNP	多型2	糖尿病	h	○
345678	SNP	多型3	糖尿病	i	○
445678	SNP	多型1	肺がん	い	○
445678	SNP	多型2	肺がん	ろ	○
445678	SNP	多型3	肺がん	は	○
456789	SNP	多型1	花粉症	j	○
456789	SNP	多型2	花粉症	k	○
456789	SNP	多型3	花粉症	l	○
456789	SNP	多型1	高血圧症	m	○
456789	SNP	多型2	高血圧症	n	○
456789	SNP	多型3	高血圧症	o	○
:	:	マイクロサテライト	多型1	不治の病	一	:	×
:	:	マイクロサテライト	多型2	不治の病	一	:	×
:	:	マイクロサテライト	多型3	不治の病	一	:	×
:	:	欠失	多型1	:	:	:	○
:	:	欠失	多型2	:	:	:	○
:	:	欠失	多型3	:	:	:	○

図4

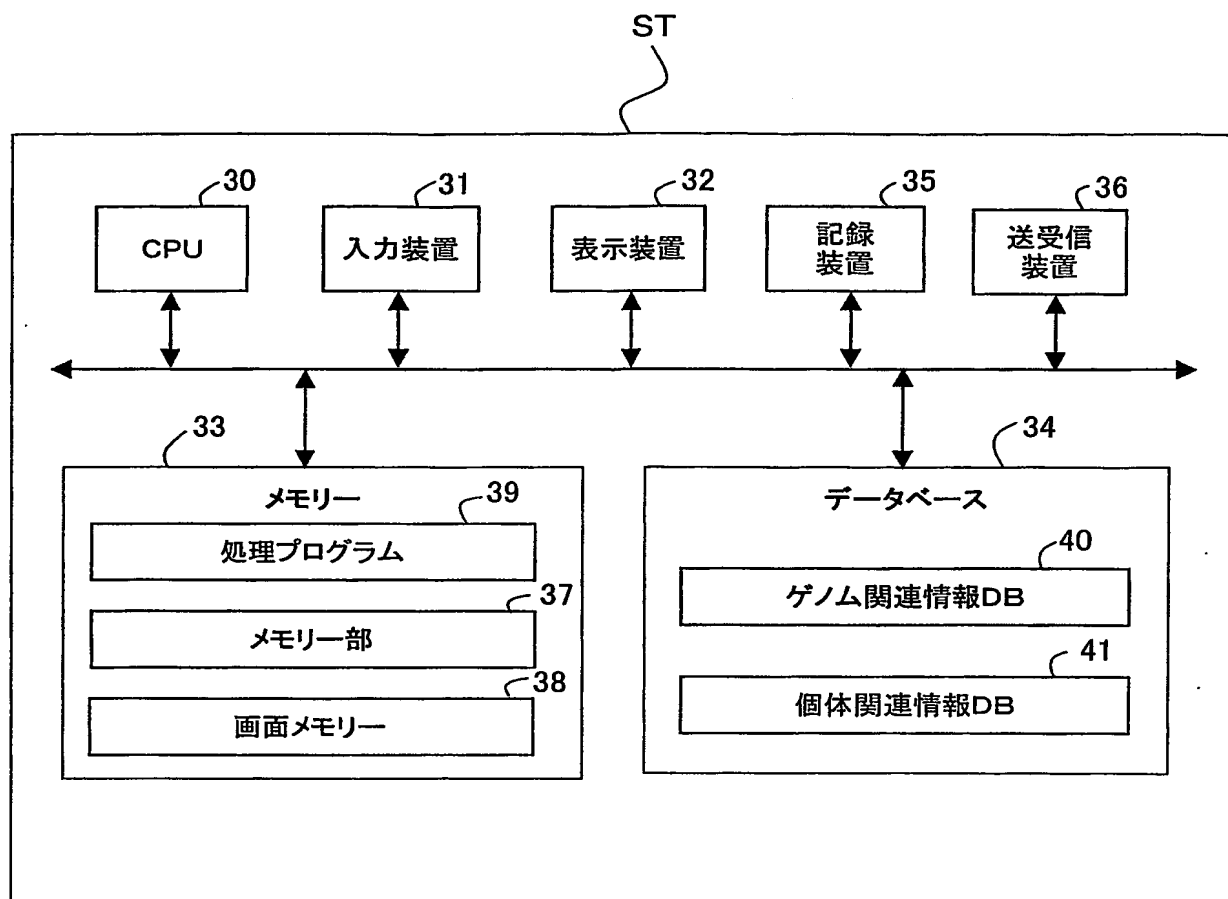


図5

多型番地	整理No. A0001	整理No. A0002	整理No. A0003
000001	多型2	多型1	多型3
000002	多型1	多型1	多型1
:	:	:	:	:
123456	多型3	多型2	多型2
:	:	:	:	:
223456	多型1	多型3	多型1
:	:	:	:	:
234567	多型1	多型1	多型1
:	:	:	:	:
334567	多型1	多型2	多型1
:	:	:	:	:
345678	多型1	多型1	多型1
:	:	:	:	:
445678	多型3	多型2	多型1
:	:	:	:	:
456789	多型1	多型1	多型1
456790	多型2	多型2	多型2
456791	多型1	多型2	多型1
456792	多型2	多型2	多型2
:	:	:	:	:

図6

個体関連情報	整理No. A0001	整理No. A0002	整理No. A0003
嗜好a	×	×	×
嗜好b	×	○	○
嗜好c	○	×	×
性質①	○	○	×
性質②	○	×	○
性質③	○	○	×
:	:	:	:	:

図7

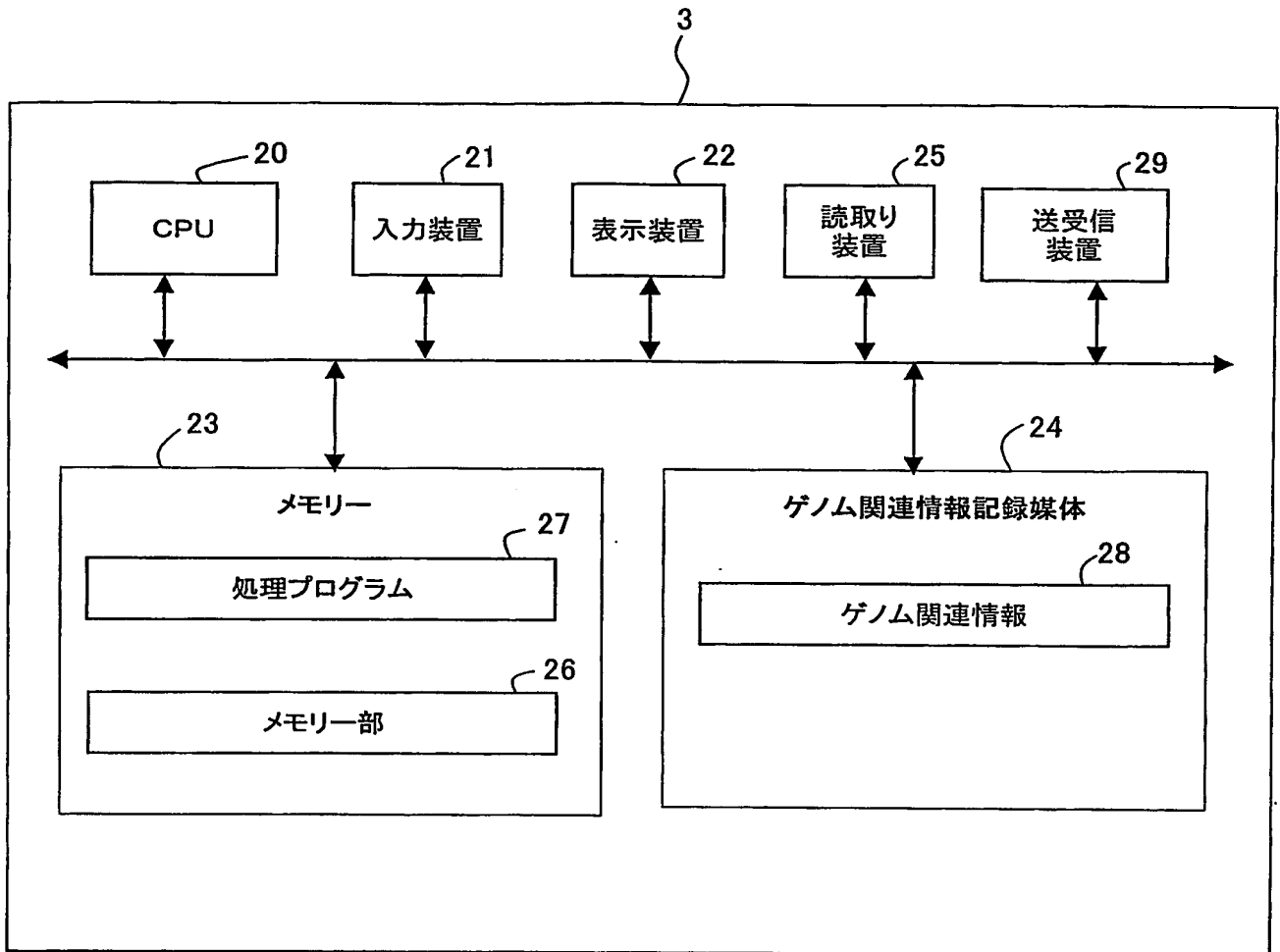


図 8

I

Gno.	生年月日
0001	***,**,**

II

多型番地	多型パターン	コメント
000001	多型1
000002	多型1
:	:	:
123456	多型2
:	:	:
223456	多型3
:	:	:
234567	多型1
:	:	:
334567	多型2
:	:	:
345678	多型1
:	:	:
445678	多型2
:	:	:
456789	多型1
456790	多型2
456791	多型2
456792	多型2
:	:	:

III

既往症
小児喘息
痛風
花粉症
胃潰瘍
アトピー
高血圧症
糖尿病

IV

特徴	記録
血液型
身長
体重
視力
走力
心理テスト
:	:
:	:
:	:
:	:
:	:

V.....

.....
(カルテ情報等)
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

図9

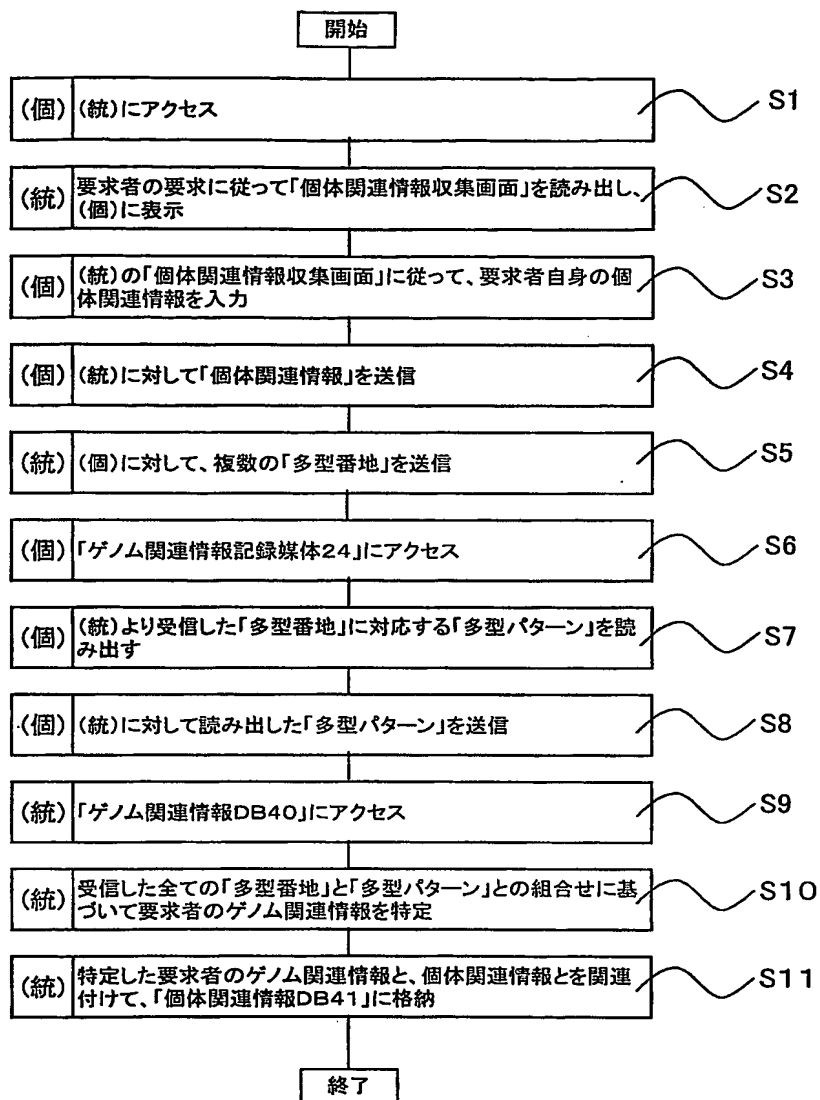


図10

<〇〇センター>

「あなたのことを教えて下さい！」（アンケート）

【アンケートに回答する前に必ずお読みください】

①あなたからご回答戴いた情報は、あなたの体質や性格等に適合した商品・サービスの提供につながるように役立てて参りますので、ご協力のほど宜しくお願い致します。

②なお、本アンケート回答を提出（送信）する際に、あなたの「Gカード」よりゲノム情報を8ヶ所※に限って、当〇〇センターにご提供戴きますので、ご承知置き下さい。※詳細は遷移画面参照
なお、あなた自身を特定し得る情報は一切戴きません。

③ご提供いただいたアンケート回答は、完全匿名化（連結不可能匿名化）された状態で当〇〇センターのデータベースに保管し、あなたの体質や性格等に適合した商品・サービスの提供に資する目的の為に役立てて参ります。アンケート回答は完全匿名化されておりますので、あなたご自身と結びつくことはありません。

④以上にご同意いただける場合は、以下のアンケートにご回答ください。

【アンケート】

[Q1]<嗜好a>

今あなたは部屋の中に独りで居ます。部屋の壁紙は赤色と青色の2種類※がありますが、どちらの色の壁紙の方が良いと思いますか？

青色の方が良いと思う方は、右のボックスをチェックして下さい。

※色見本は遷移画面参照



⋮

[Q4]<性質①>

ある人と待ち合わせの約束をしました。待ち合わせ場所までは電車に乗って行きます。

さてあなたは、万一電車が遅れても待ち合わせ時間に遅れないように、1～2本早めの電車で行った方が気が落ち着きますか？

それとも、万一のことを気にするのは無意味なので、待ち合わせ丁度に到着する電車で行った方が良いと考えますか？

どちらかといえば早めの電車で行く方に共感できるという方は、右のボックスをチェックして下さい。



⋮

●ご回答ありがとうございました。

以上でよろしければ、カードリーダーに「Gカード」を挿入の上、下記「提出」をクリック下さい。

提出

中止

図 11

多型番地	多型1	多型2	多型3	計
000001	50人	25人	25人	100人
000002	60人	20人	20人	100人
000003	40人	40人	20人	100人
000004	50人	30人	20人	100人
:	:	:	:	:

図12

多型番地	嗜好a				嗜好b				嗜好c				計
	多型1	多型2	多型3	計	多型1	多型2	多型3	計	多型1	多型2	多型3	計		
000001	45人	3人	2人	50人	..人	..人	..人	..人	..人	..人	..人	..人	..人
000002	30人	10人	10人	50人	..人	..人	..人	..人	..人	..人	..人	..人	..人
000003	21人	19人	10人	50人	..人	..人	..人	..人	..人	..人	..人	..人	..人
000004	10人	20人	20人	50人	..人	..人	..人	..人	..人	..人	..人	..人	..人
：	：	：	：	：	：	：	：	：	：	：	：	：	：	：

図13

多型番地	嗜好a				嗜好b				嗜好c			
	多型1	多型2	多型3	計	多型1	多型2	多型3	計	多型1	多型2	多型3	計	
000001	90%	12%	8%	50%	..%	..%	..%	..%	..%	..%	..%	..%
000002	50%	50%	50%	50%	..%	..%	..%	..%	..%	..%	..%	..%
000003	52.5%	47.5%	50%	50%	..%	..%	..%	..%	..%	..%	..%	..%
000004	20%	66.7%	100%	50%	..%	..%	..%	..%	..%	..%	..%	..%
∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴	∴

図14

多型番地	嗜好a			嗜好b			嗜好c		
	多型1	多型2	多型3	多型1	多型2	多型3	多型1	多型2	多型3	
000001	11.25倍	1.5倍	1.0倍	..倍	..倍	..倍	..倍	..倍	..倍
000002	1.0倍	1.0倍	1.0倍	..倍	..倍	..倍	..倍	..倍	..倍
000003	1.11倍	1.0倍	1.05倍	..倍	..倍	..倍	..倍	..倍	..倍
000004	1.0倍	3.34倍	5.0倍	..倍	..倍	..倍	..倍	..倍	..倍
:	:	:	:	:	:	:	:	:	:	:

図 15

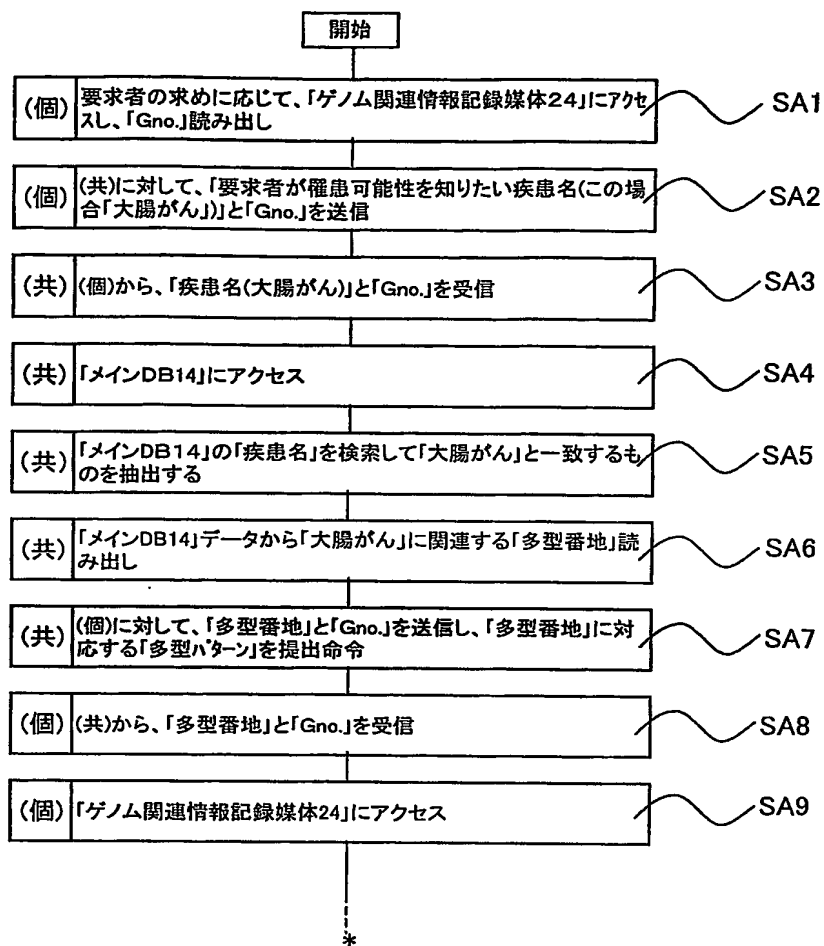


図16

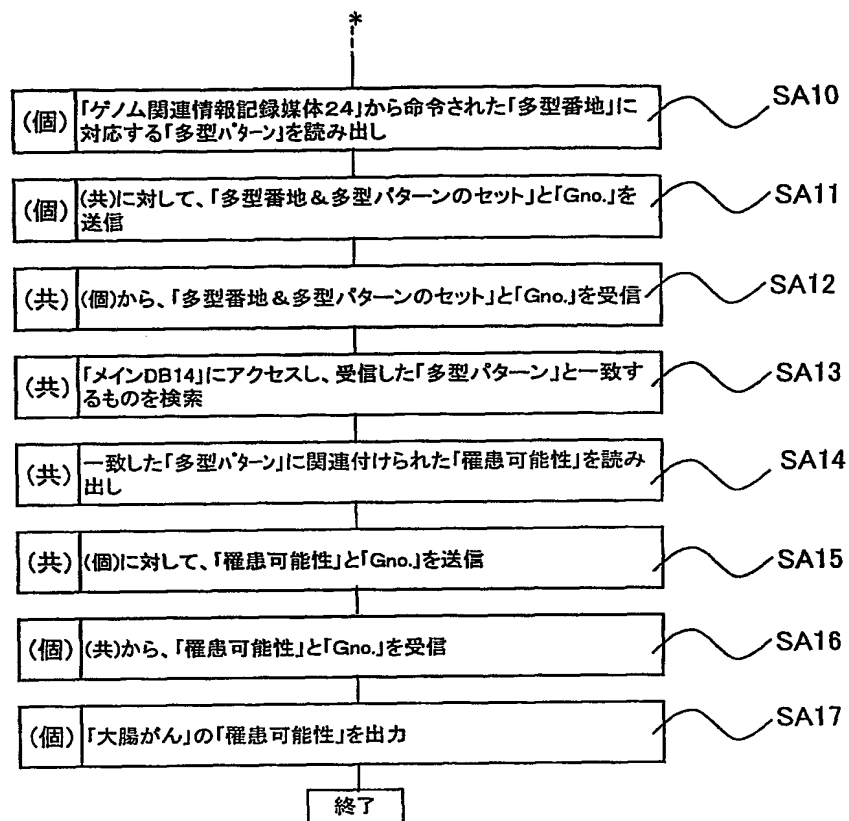


図 17

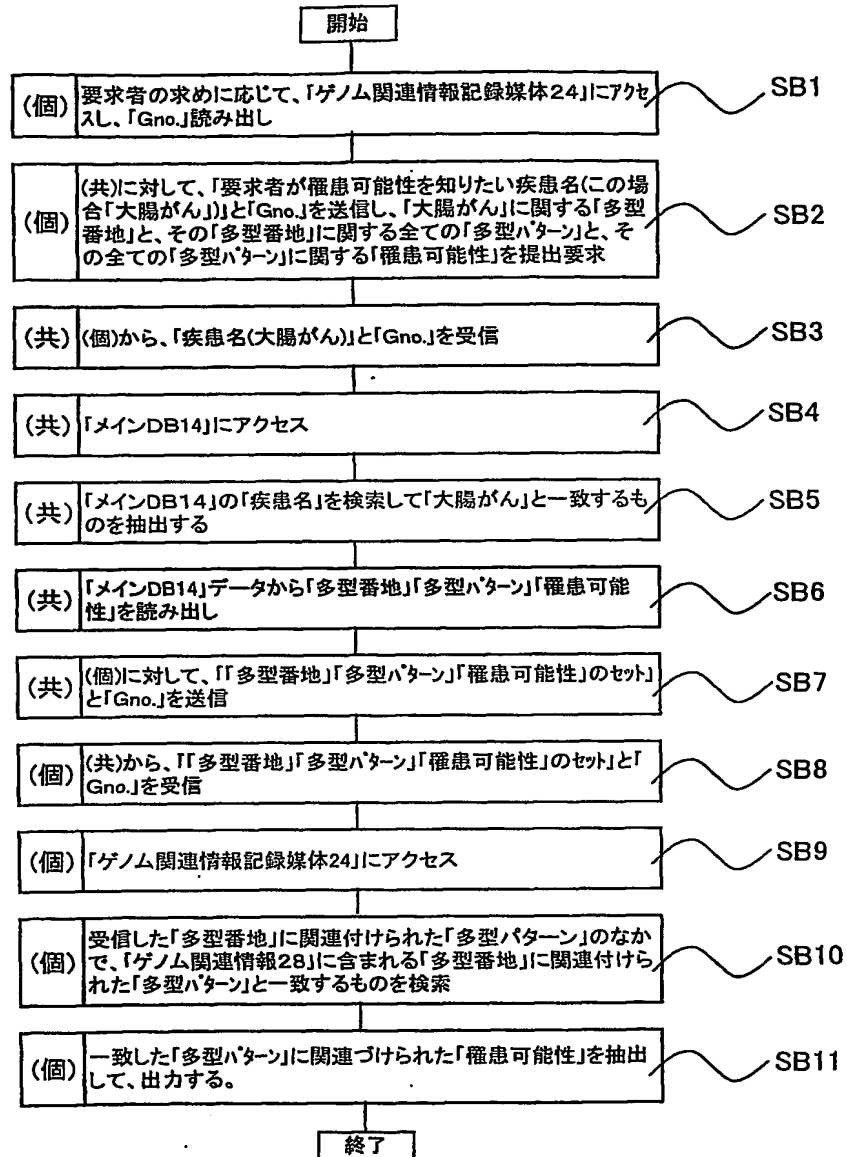
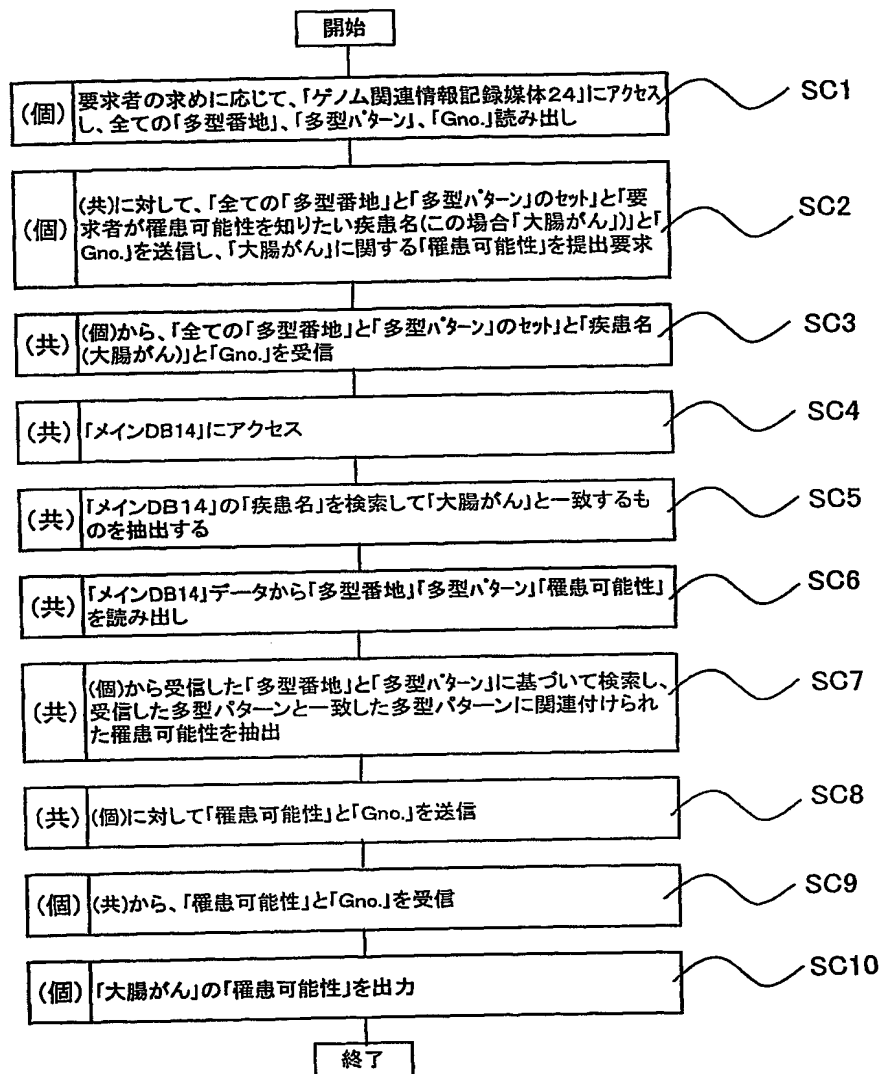


図 18



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/14733

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ G06F19/00, G06F17/60, G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G06F19/00, G06F17/60, G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST FILE (JOIS), WPI, INSPEC (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Yusuke NAKAMURA, "Lecture SNP (single nucleotide polymorphism) no Iryo Genba ni okeru Juyosei", Gendai Iryo, Vol.32, No.1, 10 January, 2000 (10.01.00), pages 204 to 210	1-6
X	Toshio FURUYA, Yasufumi MURAKAMI, "1. SNPs to Genome no Tayosei Kaiseki", Genome Kino Kenkyu Protocol, Kabushiki Kaisha Yodosha, 10 April, 2000 (10.04.00), pages 130 to 137; particularly, pages 134 to 135	1-6

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
26 January, 2004 (26.01.04)

Date of mailing of the international search report
10 February, 2004 (10.02.04)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G06F19/00, G06F17/60, G06F17/30

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G06F19/00, G06F17/60, G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2004年
 日本国登録実用新案公報 1994-2004年
 日本国実用新案登録公報 1996-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS), WPI, INSPEC (DIALOG)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	中村祐輔, レクチャーSNP (single nucleotide polymorphism) の医療現場における重要性, 現代医療, Vol. 32, No. 1, 2000. 01. 10, p. 204-210	1-6
X	古谷利夫, 村上康文, 1. SNPsとゲノムの多様性解析, ゲノム機能研究プロトコル, 羊土社, 2000. 04. 10, p. 130-137, 特に, p. 134-135	1-6

☐ C欄の続きにも文献が列举されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」口頭による開示、使用、展示等に言及する文献
 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

26. 01. 2004

国際調査報告の発送日

10. 2. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

高瀬 勤

5M

9069

電話番号 03-3581-1101 内線 3597